McGraw-Hill Electrical and Electronic Engineering Series
FREDERICK EMMONS TERMAN, *Consulting Editor*

FUNDAMENTALS OF VACUUM TUBES

McGraw-Hill Electrical and Electronic Engineering Series

FREDERICK EMMONS TERMAN, *Consulting Editor*

BAILEY AND GAULT · Alternating-current Machinery
BERANEK · Acoustics
BRUNS AND SAUNDERS · Analysis of Feedback Control Systems
CAGE · Theory and Application of Industrial Electronics
CUCCIA · Harmonics, Sidebands, and Transients in Communication Engineering
EASTMAN · Fundamentals of Vacuum Tubes
EVANS · Control-system Dynamics
FITZGERALD AND HIGGINBOTHAM · Basic Electrical Engineering
FITZGERALD AND KINGSLEY · Electric Machinery
GEPPERT · Basic Electron Tubes
GLASFORD · Fundamentals of Television Engineering
HAPPELL AND HESSELBERTH · Engineering Electronics
HARMAN · Fundamentals of Electronic Motion
HESSLER AND CAREY · Fundamentals of Electrical Engineering
HILL · Electronics in Engineering
JOHNSON · Transmission Lines and Networks
KRAUS · Antennas
KRAUS · Electromagnetics
LEPAGE · Analysis of Alternating-current Circuits
LEPAGE AND SEELY · General Network Analysis
MILLMAN AND SEELY · Electronics
MILLMAN AND TAUB · Pulse and Digital Circuits
ROGERS · Introduction to Electric Fields
RÜDENBERG · Transient Performance of Electric Power Systems
RYDER · Engineering Electronics
SEELY · Electronic Engineering
SEELY · Electron-tube Circuits
SEELY · Radio Electronics
SISKIND · Direct-current Machinery
SKILLING · Electric Transmission Lines
SKILLING · Transient Electric Currents
SPANGENBERG · Fundamentals of Electronic Devices
SPANGENBERG · Vacuum Tubes
STEVENSON · Elements of Power System Analysis
TERMAN · Electronic and Radio Engineering
TERMAN AND PETTIT · Electronic Measurements
THALER · Elements of Servomechanism Theory
THALER AND BROWN · Servomechanism Analysis
THOMPSON · Alternating-current and Transient Circuit Analysis
TRUXAL · Automatic Feedback Control System Synthesis

FUNDAMENTALS
OF
VACUUM TUBES

BY
AUSTIN V. EASTMAN, M.S.
Professor of Electrical Engineering, Head of the Department of Electrical Engineering University of Washington

THIRD EDITION

McGRAW-HILL BOOK COMPANY, INC.
NEW YORK TORONTO LONDON
1949

FUNDAMENTALS OF VACUUM TUBES

TO

UNA

PREFACE TO THE THIRD EDITION

The third edition of "Fundamentals of Vacuum Tubes" follows the general pattern of the second edition but with extensive revisions of the text material. The most extensive changes are in Chapter 9, Audio-frequency Amplifiers, and in the chapter on modulators and demodulators which has been broken down into three chapters in the third edition.

In Chapter 9 the material on resistance-coupled amplifiers has been almost entirely rewritten. The new edition treats the gain of the amplifier as a vector and thus takes into account changes in phase as well as changes in amplitude. This provides a much better foundation for later work, especially in the treatment of feed-back amplifiers. Video amplifiers are also more fully covered, including some analysis of low-frequency and high-frequency compensating circuits.

The power-series method of analysis has been removed from Chapter 9 and simpler methods are used in the development of the equations for the power amplifier. The power-series method of analysis is then fully covered in a new chapter, Chapter 12. This chapter may be omitted by those who are not interested in a general mathematical approach to the problems of vacuum-tube performance and who are willing to skip over some of the mathematical treatments of modulation and demodulation in Chapters 13 and 14.

Modulation and demodulation are covered in separate chapters, instead of being studied together, as in the second edition.

As in the earlier editions, extensive footnote references are included to permit the student who so desires to study further details in the literature. The intent has been to make complete and accurate the material presented in this book but to leave out minor details and exceptions to normal operating conditions, relying upon the footnote references to meet the needs of those who require more extensive knowledge of any specific phenomenon.

Some thought was given to removing references which were more than a few years old, but it seemed to the author that even the older references might prove of value to many students.

Therefore all but a few of those which appear in the second edition also appear in this third edition.

Early plans for the third edition included one or more chapters on ultra-high-frequency tubes and their circuits. Further consideration led to the conclusion that either the book would then become undesirably long or that the treatment of these tubes and circuits would have to be seriously curtailed. It was therefore decided to omit all discussion of ultra-high-frequency tubes and techniques. Many excellent texts are now available dealing with the u-h-f field alone, and thus may be used to supplement the fundamental studies presented in this book.

Most of the problems of the second edition have been retained but a number of new ones have been added. A frequent argument for providing all new problems when revising a book is that students may otherwise secure answers from those who have used the book before. It is the author's belief that junior and senior engineering students should be sufficiently serious about their studies to work out their own problems regardless of whether or not answers are available. Thus it seems desirable to retain problems which have proved satisfactory in the past. For this same reason the author makes answer books available to all his students and every effort is being made to have an answer book for the third edition available at the time of publication.

The mks rationalized units have become widely accepted in electrical-engineering work. Consequently they have been generally employed in this third edition, although cgs units have been retained in a very few places where the size of the unit or general usage seemed to make their use desirable. The mks units should be assumed unless others are specified.

AUSTIN V. EASTMAN

SEATTLE, WASH.
July, 1949

PREFACE TO THE FIRST EDITION

It has been the author's intent in writing "Fundamentals of Vacuum Tubes" to combine in a single text the basic theory underlying the operation of all types of modern vacuum tubes, both radio and industrial, together with their more common applications. In so doing an effort has been made to avoid either writing a text that is almost purely descriptive or presenting such a wealth of mathematical material as to make the work unattractive to a large group of potential readers who desire only a basic working knowledge of vacuum tubes. Although books going to either extreme are unquestionably valuable to certain groups, it is the author's belief that these needs are already well supplied with excellent texts.

A knowledge of the basic laws of direct and alternating currents is assumed together with mathematics up to and including the simple calculus. However, the use of the calculus is such that anyone unfamiliar with this branch of mathematics may still follow the essential points presented in this text. Therefore, although the book is designed especially for senior electrical students in an engineering college, it is hoped that it may prove of benefit to other classes of students as well as to practicing engineers.

This book deals with the laws governing the operation of the vacuum tube itself instead of presenting a large number of circuits in which the tube may be used. Consequently circuits have been selected and described only for the purpose of illustrating the operation of the tube and no attempt has been made to present a complete discussion of such individual applications as television, radio broadcasting, etc., although many footnotes are provided giving reference to more extensive treatises on the various circuits. It is believed that this method of treatment will enable the reader to obtain a thorough grounding in vacuum-tube fundamentals with a minimum expenditure of effort and will provide a groundwork from which he can readily familiarize himself with the details of any special application by a little

additional study For this reason the text should prove of value to anyone interested in the electronic art whether in the industrial or the communication field.

It has been the author's experience in teaching vacuum-tube courses at the University of Washington that often an otherwise excellent and convincing mathematical analysis of some phenomenon will be entirely incomprehensible to the student because of his inability to see the objective toward which the study is directed. To avoid this pitfall the author has preceded many of the mathematical analyses in this text with a thorough verbal description of the phenomena The student can then follow the mathematical development with a mental picture of the phenomena and so interpret the various steps in terms of the physical reactions that they illustrate Thus, the mathematical treatises become the tools by which the more complex phenomena can be understood rather than a thing apart, as they too often become.

The more fundamental equations in the text have been marked with an asterisk (*) following the equation number to assist those wishing to use the book as a reference It cannot be overemphasized, however, that great care must be exercised to see that the equation selected actually covers the situations being studied. A careful perusal of the subject matter of the text just preceding the equation should safeguard against such errors.

The author wishes to express his appreciation to his wife, to whom this book is dedicated, for her unfailing encouragement throughout the preparation of the manuscript, to Professor Frederick E. Terman and Mr. B J. Thomson for valuable suggestions and information; to the General Electric Co. for suggestions as well as for many illustrations, and to the Westinghouse Electric & Mfg. Co., R. C. A. Mfg Co., and Raytheon Production Corp. for illustrations.

AUSTIN V. EASTMAN

SEATTLE, WASH.
April, 1937

CONTENTS

Preface to the Third Edition v
Preface to the First Edition vii
List of Tables xxi

PART I

BASIC CONCEPTS

Introduction to Part I 3

Chapter 1

ELECTRONIC EMISSION

Edison Effect 5
Electron Theory 5
Emission of Electrons 7
Thermionic Emission 7
Oxide-coated Cathodes 9
Thoriated-tungsten Cathodes 10
Carbonization of Thoriated Filaments 10
Measurement of Emission 11
A-C Heated Cathodes 12
Laws Governing Emission 13
High-field Emission 17
Photoelectric Emission 18
Theory of Photoelectric Emission 18
Threshold Frequency 20
Photoelectric Emission from Composite Surfaces 20
Effect of Light Intensity 20
Spectral Selectivity 21

Chapter 2

CONSTRUCTION OF VACUUM TUBES

General 23
Cathodes 27
Grids 28
Anodes 29
Methods of Cooling 29

Evacuation of Vacuum Tubes 33
✓Getter 34

✓CHAPTER 3

HIGH-VACUUM TUBES

1 DIODES

[illegible] Diodes 37
[illegible] . . . 39
[illegible] 41
Space-charge Equation 43
Energy Loss on Plate 45
A C Plate Resistance 47
Effect of Gas . 48
Rectifier Action 50

2 TRIODES

Static Characteristic Curves 51
Potential Distribution 54
Effectiveness of Grid in Controlling Plate-current Flow 55
Amplification Factor . 57
Amplification of Alternating Currents . . . 58
Analysis of an Amplifier . . . 60
Phase Relations in an Amplifier 61
Vacuum-tube Coefficients . . . 65
Grid Current . . 68

3 MULTIGRID TUBES

Tetrodes Screen-grid Tubes 70
Tetrode Static Characteristic Curves 72
Potential Distribution 75
Variation in μ, r_p, and g_m 76
Pentodes 77
Pentode Characteristic Curves 78
Potential Distribution 79
Remote Cutoff Tubes 82
Beam Power Tube 82
Potential Distribution Curves 85
Multigrid Tubes as Amplifiers 86

4 MISCELLANEOUS HIGH-VACUUM TUBES

Pentagrid Converters and Mixers 88
Space-charge-grid Tubes 89
Acorn Tubes 89
Lighthouse Tubes 90

CHAPTER 4

GAS-FILLED TUBES

Current Flow through Gases 93

Glow Discharges 97
Arc Discharges 98
Arc Back 99
Cathode Disintegration by Positive-ion Bombardment 99

1. DIODES

Hot-cathode, Mercury-vapor Tubes 100
Mechanism of Current Flow 101
Maximum Permissible Current Flow 103
Heat-shielded Cathode 104
Temperature Limits 105
Mercury-arc Tubes. Construction Details 106
Mechanism of Current Flow 108
Arc Back 109
Mercury-arc vs. Mercury-vapor Tubes 110
Ignitron 112
Comparison of Ignitron with Mercury-arc and Mercury-vapor Tubes 113
Tungar Tubes 115
Mechanism of Current Flow 116
• Cold-cathode Tubes 117

2. TRIODES

• Thyratrons 120
Potential Distribution Curve 121
Type of Grid Structure Required 122
Thyratron Characteristic Curves 123
Use of Gases Other Than Mercury Vapor 124
Comparison of the Thyratron with the High-vacuum Tube 125
Grid Control of Mercury-arc Tubes 126
Grid Control of Ignitrons 126
Grid-controlled Cold-cathode Tubes 126

3. SCREEN-GRID TUBES

Screen-grid Tubes 129

CHAPTER 5

PHOTOELECTRIC TUBES

Units Used in Light Measurements 132

1. PHOTOEMISSIVE TUBES

Characteristic Curves. High-vacuum Tubes 133
Characteristic Curves. Gas-filled Tubes 135
Color Response of Phototubes 136
Phototube Circuits 137

2. PHOTOCONDUCTIVE CELLS

Selenium Cells 138
Other Types of Cells 140
Circuits Suitable for Photoconductive Cells 140

3. PHOTOVOLTAIC CELLS

Iron Selenide Cell 142
Equivalent Circuit of Photovoltaic Devices 142
Circuits Suitable for Use with Photovoltaic Cells 143

CHAPTER 6

SPECIAL TYPES OF TUBES

Cathode ray Tubes 145
Oscilloscopes 149
Electron-ray Tube 150
Electron Multiplier 151
Special Tubes Used for Television 153
X-ray Tubes 158

PART II

APPLICATIONS AND CIRCUITS

Introduction to Part II 163

CHAPTER 7

RECTIFIERS

Single-phase, Half-wave Rectifier 164
Single-phase, Full-wave Rectifier 165
Single-phase, Bridge Rectifier 167
Polyphase Rectifier Circuits 168
Three-phase, Half-wave Rectifier 168
Three phase, Full-wave Rectifier 171
Three-phase, Bridge Rectifier 174
Choice of Tubes and Circuits 178
Filters 181
Current Flow in Rectifiers When Filtered 183
Tuned-circuit Filters 186
L-section Filter Design 187
Minimum Capacitance for Filter 189
Minimum Inductance for Filter 190
Design of Choke 191
Selection of Tube 193
Complete Rectifier Design 196
Pi section Filters 199
Operating Characteristics 201
Rectifiers Using R-F Power 202
Voltage multiplying Circuits 204
Special Circuits for Use with Ignitrons 206
Special Igniter Circuits to Ensure Firing 208
Other Types of Firing Circuits 209
Special Circuits for Mercury-arc Rectifiers 211
Operating Characteristics 213
Controlled Rectifiers 214

CHAPTER 8

THE VACUUM TUBE AS A CONTROL DEVICE

1. HIGH-VACUUM TUBES

High-vacuum Tubes........ 217
A-C Operation........ 218
Operation of Mechanical Relays........ 221
A-C Operation of Relays........ 223
Voltage-regulated Power Supplies........ 224

2. GAS-FILLED TUBES

Gas-filled Tubes........ 227
D-C Circuits for Use with Thyratrons........ 227
A-C Operation........ 229
Plate-current Control by Grid Phase Shift........ 231
Wave Shapes........ 233
Circuits with Alternating Current through the Load........ 234
Controlled Rectifier........ 236
Separately-excited Inverters........ 240
Self-excited Inverters........ 243
Phase-shift Control of Ignitrons........ 246
Applications of Thyratrons and Ignitrons........ 247
Operation of Cold-cathode, Grid-controlled Tubes........ 249

3. PHOTOELECTRIC TUBES

Photoelectric Tubes........ 251
Phototube Control of Thyratrons........ 252
Photovoltaic Cells as Control Devices........ 254

CHAPTER 9

AUDIO-FREQUENCY AMPLIFIERS

Classes of Amplifiers........ 257
Distortion........ 258
Decibels........ 259
Input Admittance of a Vacuum Tube........ 260

1. VOLTAGE AMPLIFIERS

Resistance-coupled Amplifiers........ 264
Equivalent Circuits of Resistance-coupled Amplifiers........ 265
Gain of a Resistance-coupled Amplifier........ 268
Frequency Response of Resistance-coupled Amplifiers........ 272
Universal Amplification Curves........ 274
Characteristics of Suitable Tubes........ 275
Circuit Elements in Resistance-coupled Amplifiers........ 276
Practical Circuits........ 278
Effect of Screen-grid By-pass Condenser........ 278
Effect of Cathode By-pass Condenser........ 284
Minimum Frequency Response of Pentode Amplifiers........ 284
Multistage Amplifiers........ 285

Low-frequency Compensation 286
High-frequency Compensation 290
Transformer-coupled Voltage Amplifiers 294
Gain of a Transformer coupled Amplifier 297
Frequency Response of a Transformer-coupled Amplifier 299
Resistance Loading in Transformer-coupled Amplifiers 302
Impedance-coupled Amplifiers 303
Cathode-ray Oscilloscopes 303
D C Amplifiers 307
Amplifiers Controlled by Phototubes 310
Volume control Methods 311
Sources of Noise in Amplifiers 314
Hum 315
Microphonic Noise 316

2 POWER AMPLIFIERS

Class A Power Amplifiers 317
Conditions for Maximum Power Output Resistance Load 319
Theoretical Determination of Limit Resistance to be Used 323
Effect on Distortion of Changes in Load Resistance 327
Graphical Determination of Power Output Triodes 328
Graphical Determination of Power Output Pentodes 334
Effect of Reactive Loads 337
Plate Efficiency Triode Tubes 338
Plate Efficiency Pentode and Beam Tubes 340
Push-pull Amplifiers 341
Class A Push-pull Amplifier Performance 343
Composite Characteristic Curves 344
Theoretical Determination of Correct Load Resistance 346
Class B Push-pull Amplifiers 348
Effect of Driving Grid Positive 350
Class AB Push-pull Amplifiers 352
Plate Efficiency of Push-pull Amplifiers 354
Phase-inverting Tube to Drive a Push-pull Amplifier 355
Frequency Distortion in Power Amplifiers 357

3 FEED BACK AMPLIFIERS

Analysis of Feed-back Amplifiers 361
Effect of Feedback on the Frequency Response of an Amplifier 364
Current Feedback 367
Feed-back Circuits 368
Cathode-follower Amplifier 371
Regeneration 375

CHAPTER 10

✓ RADIO-FREQUENCY AMPLIFIERS

Radio-frequency Amplifiers 382

1. VOLTAGE AMPLIFIERS

Impedance-coupled Amplifiers.................................... 383
Transformer-coupled Amplifiers.................................. 386
Band-pass Filters... 388
Multistage Amplifiers... 394

2. POWER AMPLIFIERS

Power Amplifiers.. 394
Field of Use of Class B and Class C Amplifiers.................. 396
Typical Circuits for Class B and Class C Amplifiers............. 397
Wave Shapes in Class C Amplifiers. Series Feed.................. 398
Wave Shapes in Class C Amplifiers. Parallel Feed................ 401
Power Required to Drive a Class C Amplifier..................... 403
Design of Class C Amplifiers. Exact Method...................... 404
Approximate Design of Class C Amplifiers. Triodes............... 405
Design of Class C Amplifiers Using Pentodes and Tetrodes........ 413
Design of the Tank Circuit to Produce Correct Plate Voltage..... 414
Energy Storage in Tank Circuits................................. 416
Design of the Tank Circuit to Secure Sufficient Energy Storage.. 417
Significance of Q in the Design of Class C Amplifiers......... 419
Approximate Solution for the Tank Circuit of Class C Amplifiers. 421
Sources of Power for Class C Amplifiers......................... 422
Neutralization.. 423
Bridge-type Neutralizing Circuits............................... 425
Adjusting Class C Amplifiers.................................... 427
Frequency Doublers.. 429
Class B Linear Amplifiers....................................... 430
Doherty High-efficiency Amplifier............................... 432
Applications of Class B Amplifiers.............................. 435
Grounded-grid Amplifier... 436
Cathode-coupled Amplifier....................................... 438

CHAPTER 11

OSCILLATORS

Oscillators... 441
Oscillator Circuits... 442
General Equations of an Oscillatory Circuit..................... 445
Applications of the General Equations........................... 452
Power Oscillators... 454
Frequency Stability... 457
Methods of Increasing Frequency Stability....................... 459
Grid-condenser Design... 461
R-C Oscillator.. 462
Phase-shift Oscillator.. 467
Crystal Oscillators... 467
Characteristics of Quartz Crystals.............................. 469
Zero-temperature-coefficient Crystals........................... 470
Equivalent Circuit of the Crystal............................... 473

Bridge-stabilized Oscillator . . . 475
Use of Crystal Oscillators . . . 476
Frequency Control with Resonant Lines . . . 476
Dynatron . . . 477
Negative Transconductance Oscillator . . . 479
Beat frequency Oscillator . . . 480
Multivibrators . . . 481
Controlled Multivibrators . . . 482
Parasitic Oscillations . . . 484

CHAPTER 12

POWER-SERIES ANALYSIS OF VACUUM-TUBE PERFORMANCE

Power-series Expansion of the Plate Current Equation . . . 488
Amplitude Distortion . . . 495
Power Series Expansion of the Plate Current Diodes . . . 497
Effect of Load Impedances Which are a Function of Frequency . . . 500
Effect of Applying Two Signals of Different Frequencies . . . 502
Determination of r_p', μ_b', and μ_c' . . . 504
Diodes . . . 505
Pentods and Beam Tubes . . . 505
Components Produced by Distortion . . . 505
Applications of the Power-series Analysis . . . 506

CHAPTER 13

✓ MODULATORS

Modulation . . . 508
Uses of Modulation . . . 510

1. AMPLITUDE MODULATION

Amplitude Modulation . . . 510
Side Bands . . . 512
Power in a Modulated Wave . . . 513
Linear Modulators . . . 514
Linear Plate Modulation . . . 514
Sources of Modulating Voltage for Plate-modulated Amplifiers . . . 516
Plate Modulation with Class B Modulators . . . 518
Distortion in Plate modulated Class C Amplifiers . . . 519
Plate Modulation of Tetrode and Pentode Tubes . . . 520
Grid Modulation . . . 521
Distortion in Grid modulated Amplifiers . . . 523
Suppressor-grid Modulation of Pentodes . . . 524
Use of Cathode-ray Tube in Checking Modulation . . . 525
Square-law Modulation . . . 526
Analysis of the Performance of a Square-law Modulator . . . 529
Components of Square-law Modulation . . . 531
Cross Modulation . . . 531
Balanced Modulator . . . 533
Single Side band Transmission . . . 534

2. FREQUENCY AND PHASE MODULATION

Frequency and Phase Modulation 536
Analysis of Frequency and Phase Modulation 536
Band Width 540
Analysis by Vector Diagrams 541
Noise and Interference Reduction 546
Methods of Producing Frequency Modulation 548
Balanced Modulator System of Frequency Modulation 549
Reactance-tube Modulator 553
Synchronized Frequency Modulation 556

3. PULSE MODULATION

Pulse Modulation 557

CHAPTER 14

✓ DEMODULATORS

Demodulation 563

1. DEMODULATION OF AMPLITUDE-MODULATED WAVES

Large-signal Diode Detectors 563
Large-signal Triode Detectors 565
Analysis of the Peak Linear Detector 566
Distortion in Peak Linear Detectors 570
Infinite-impedance Detector 573
Square-law (Weak-signal) Detectors 574
Performance of a Square-law Detector 575
Components of Square-law Demodulation 577
Grid-leak Detector 577
Heterodyne Detection 578
Principle of Operation of Superheterodyne Receivers 582
Pentagrid Converter Tubes in Superheterodyne Receivers 583
Pentagrid Mixer Tubes in Superheterodyne Receivers 584
Autodyne, or Oscillating Detector 586
Regenerative Detectors 587
Automatic Volume Control 587
Vacuum-tube Voltmeter 591
Characteristics of Vacuum-tube Voltmeters 594

2. DEMODULATION OF FREQUENCY-MODULATED WAVES

Amplitude Limiter 597

3. DEMODULATION OF PULSE-MODULATED WAVES

APPENDIX A

DEFINITIONS AND NOMENCLATURE

Definitions 601
Symbols 607

APPENDIX B

Fourier Analysis of a Repeating Function 613

APPENDIX C

Application of the Fourier Analysis to an Analytical Solution of Repeating Functions, the Equations of Which Are Known over Short Intervals 617

APPENDIX D

Methods of Evaluating the Numerical Values in Table 7-1 618

APPENDIX E

Relative Polarities of Currents and Voltages in a Vacuum-tube Amplifier 626

AUTHOR INDEX 631

SUBJECT INDEX 635

LIST OF TABLES

TABLE

1-1. Values of b_0, ϕ, and A for Various Materials.. 14

7-1. Rectifier Design Data for Typical Rectifier Circuits............ 195

9-1. Results of Fourier Analysis of the Curve of Fig. 9-46 330

13-1. Example of Plate-dissipation Requirements of the Modulator and Class C Amplifier in the Circuit of Fig. 13-9.................... 518

14-1. Substitutions to be made in Applying the Analysis of Class A Amplifiers to Peak Linear Detectors............................ 570

A-1. Vacuum-tube symbols.. 609

A-2. Vacuum-tube symbols 610

A-3. Electrical and Vacuum-tube Circuit Elements.................. 610

A-4. Abbreviations... 612

B-1. Fourier Analysis of Curve of Fig. 9-46......................... 615

PART I

BASIC CONCEPTS

INTRODUCTION TO PART I

The field of electrical engineering has grown so wide in scope that it is commonly subdivided into a number of smaller divisions. Of these divisions the two largest seem to be communication and power. In fact many engineers have used these terms somewhat loosely to include all electrical engineering in these two divisions by classifying as power all fields not definitely of a communication nature. It is in this sense that the terms *communication* and *communication engineering* and the terms *power* and *power engineering* are used in this book, although in some cases the term *industrial* is used instead of power.

In its initial stages of development and application the vacuum tube was definitely a communication device and not a power device. This was due in large part to the low current-carrying capacity of the early tubes and, perhaps, in some measure to the reluctance of the power engineer to trust a device that looked as fragile as the glass-enclosed vacuum tube. The development of gas-filled tubes and, to a lesser extent, the frequent use of metal envelopes in place of glass have largely overcome these barriers, and today it is important that the power engineer be as familiar with the operation of vacuum tubes as the communication engineer. On the other hand, many applications of the vacuum tube that are of major importance to the communication engineer are of slight interest to the power engineer and, conversely, many power applications are of little interest to communication engineers.

This book has been divided into two parts in an attempt to meet these various needs. The material of Part I is intended to lay the foundation work for all who are interested in vacuum tubes from any point of view, since basic concepts are the same regardless of whether the tube is to be used for communication or for power purposes. In some cases simple circuits and applications are presented in this part to make clear the purpose of certain tube constructions and designs or because the applications are too elementary or too unimportant to warrant separate treatment in Part II.

Part II is devoted to a study of the basic circuits and applications of the vacuum tube. The first two chapters (Chaps. 7 and and 8) describe those basic applications and circuits which are probably of primary interest to the power engineer; the remaining six present those which are of primary interest to the communication engineer. Obviously it is impossible to make an absolute segregation along any such lines, therefore, power engineers may well find much of interest in the later chapters, especially in Chap. 9, whereas communication engineers will find certain portions of the first two chapters of value to them, especially the material in Chap. 7.

It may be desirable in certain cases to go directly into the material of Part II after completing only the first four chapters of Part I, thus omitting photosensitive devices and special tubes. This may readily be done, since very little of the material in Part II is based on the last two chapters of Part I, these two chapters being largely complete in themselves.

It is suggested that students secure a copy of one of the tube manuals issued by leading tube manufacturers, to be used for reference purposes. These manuals give tube constants, rated voltages and currents, and other valuable information.

CHAPTER 1

ELECTRONIC EMISSION

Thomas Alva Ed

24|1|

Edison Effect. The vacuum tube may well be considered as having had its birth in the now common electric lamp. In 1883 Thomas A. Edison was busily engaged in endeavoring to produce a better, longer lived electric lamp than he had hitherto been able to produce. One of the experiments that he conducted for this purpose was the placing of a metallic plate inside the evacuated bulb. In the process of his experiment he connected a galvanometer between this plate and the filament, as in Fig. 1-1, and observed a small current flowing. He also observed that, when the wire from the plate was connected to the positive terminal of the filament battery, the indication of the galvanometer was greater than when the return was made to the negative end. This phenomenon is known as the *Edison effect.*

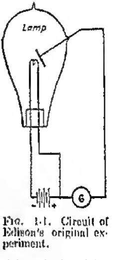


FIG. 1-1. Circuit of Edison's original experiment.

Electron Theory. It was several years after the discovery of the Edison effect before scientists were able to advance a theory satisfactorily explaining this phenomenon. About 1900 Sir J. J. Thomson, as the culmination of several years' work, advanced the *electron theory*. Briefly, this theory considers all matter to contain very small particles of electricity called *electrons*. These electrons are apparently identical, having a mass of $(9.1055 \pm 0.0012) \times 10^{-28}$ g and an electric charge of $(1.60199 \pm 0.00016) \times 10^{-19}$ coulomb.

Later studies indicate that matter is actually made up of atoms each of which consists of a positive nucleus surrounded by one or more electrons traveling in well-defined orbits. Each basic element has a fixed number of electrons per atom, the number of electrons being equal to the atomic number of the element. Thus

hydrogen has one electron, helium two, etc. Elements with higher atomic numbers have two or more orbits in which electrons revolve, whereas hydrogen and helium have but one. Under certain conditions of excitation it is possible for an electron to be thrown from its normal orbit into one farther out from the nucleus. Since it requires energy to move the electron to the new orbit, the various orbits represent increasing energy levels with the lowest level adjacent to the nucleus. An electron which has received additional energy sufficient to cause it to jump from its normal orbit to one of higher energy level is unstable and will usually drop back into its normal position. In so doing it must, of course, lose the additional energy which it had previously gained, and this energy is emitted in the form of radiation, including visible light. Thus gases may give off light under certain conditions, a principle which is made use of in modern tubular lamps filled with neon or other inert gases. This phenomenon is more fully described in Chap. 4.

Some of the electrons in the outer orbit of an atom are not rigidly bound to the atom but are relatively free to travel around more or less independently from one atom to another. These are known as *free electrons*. Free electrons are most likely to be present when the number of electrons in the outer orbit is less than the maximum number which that orbit can accommodate. Thus the inert gases each have a maximum number of electrons in their outer orbits, and not only have they no free electrons but they are chemically inactive. Certain materials, such as copper, aluminum, nickel, and iron, have a relatively large number of free electrons and are called *conductors*. Others, such as glass, dry wood, silk, and porcelain, have relatively few free electrons and are known as *nonconductors* or *insulators* (although certain insulators, such as glass and porcelain, become conductors at higher temperatures). While there is no absolute dividing line between these two classifications, most materials fall fairly definitely into one group or the other.

The atoms and electrons in any material are ordinarily in rapid vibratory motion, the velocity of their motion being a function of temperature. In solids they are constrained within the limits of the material by a *potential barrier* which maintains the solid in a given shape. In liquids there is much less surface restraint, while in gases there is none.

Whenever an electron has been removed from an atom, it leaves

that atom with a positive charge equal in magnitude to the charge on the electron. Such atoms are known as *positive ions*. They are many times heavier than an electron and vary widely in weight and sometimes in charge, depending upon the material that they constitute and the method of ionization.

If the terminals of a battery or other source of emf be connected to two points of a material containing free electrons, there will be a general, though relatively slow, movement of free electrons from the negative terminal toward the positive (Fig. 1-2). The rate at which these electrons pass a given point in the circuit determines the magnitude of the *electric current*. If sufficient emf is applied to cause 6.25×10^{18} electrons to pass any one point in the circuit during the time of 1 sec, a current of 1 amp is said to be flowing.

It should be noted at this point that the flow of electrons is actually just opposite to the usually assumed direction of flow of a positive current. It really makes little difference as to which direction of current flow is assumed positive in so far as studying electrical phenomena is concerned; so in this text, as in most others, the positive direction of current flow will be assumed from a positive to a negative potential, or opposite to the direction of flow of electrons.

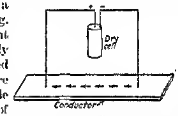


FIG. 1-2. Illustrating the flow of electrons through a conductor (conventional *positive* direction of current flow is opposite to that shown for electronic flow).

Emission of Electrons. It is possible to increase the energy of the free electrons in a conductor until they are able to pass through the potential barrier into space. They are then said to have been emitted. The four most common methods of producing such emission, listed in their order of importance, are:

1. Thermionic emission.
2. Photoelectric emission.
3. Secondary emission.
4. High-field emission.

Thermionic Emission. As previously stated, the velocity of the electrons and atoms, as they move about within the confines of the material that they comprise, is dependent upon the temperature. At a temperature of absolute zero all molecular activity is supposed to cease. As the temperature is increased, the activity

of the electrons and atoms increases until a point is reached where the electrons have sufficient velocity to enable them to break through the potential barrier of the material. This evaporation of electrons from the body of a solid at high temperatures is known as *thermionic emission*. The emission or evaporation of electrons takes place at a lower temperature than does that of the atoms; the mass of the electron being smaller, it reaches the higher velocities necessary for evaporation at lower temperatures than does the heavier atom. When the temperature becomes high enough for the atoms to evaporate, the material or solid that they compose rapidly disintegrates.

It is possible to compute the velocity that an electron must attain in order to be emitted from any material. This velocity varies with the type of material used. Tungsten, for example, has a work function of 4.53 volts; *i.e.*, the electron must possess kinetic energy in the amount of 4.53 joules/coulomb in order to break through the potential barrier, or, since an electron has a charge of 1.60×10^{-19} coulomb, 7.25×10^{-19} joule is required for the emission of an electron from tungsten. The velocity may, therefore, be determined from the following equation:

$$\text{Kinetic energy} = \frac{mv^2}{2} = \phi e \qquad (1\text{-}1)$$

where $m = 9.1 \times 10^{-31}$ kg
$e = 1.60 \times 10^{-19}$ coulomb
$\phi = 4.53$ volts (work function for tungsten)
This yields a velocity of $v = 1.26 \times 10^6$ m/sec.

Other materials have different values of work function, some higher, some lower. One might suppose that the material having the lowest work function would be the most suitable as a source of electron emission; but, as the emitting surface must possess certain mechanical features, this is not necessarily true. Two outstanding materials having lower work functions than tungsten are thorium (3.35) and calcium (2.24), but both vaporize at temperatures that are too low to produce satisfactory emission.

Pure tungsten filaments, similar to those in incandescent lamps, were used as a source of emission in the early vacuum tubes. These operated at a temperature of approximately 2500 to 2600°K[1]

[1] The K stands for Kelvin. To convert from degrees centigrade to degrees Kelvin requires the addition of 273°, *e.g.*, 2227°C = 2500°K.

and required a current for heating purposes varying from 1 amp at 5 volts for receiving tubes to as high as 52 amp at 22 volts for a 10-kw power tube. The comparatively large amounts of power required by the early receiving tubes constituted a serious drawback to the development of multitube, sensitive receivers, especially as it was necessary to supply this power from storage batteries or dry cells. A five-tube set required 25 watts, and a 75-amp-hr storage battery would supply it for only 15 hr.

The emission efficiency (ratio of emission current to heating power) of a tungsten cathode is approximately 5 to 10 ma/watt at normal operating temperature. Operation at a higher temperature will, of course, increase the emission efficiency but will shorten the life, whereas a lower temperature will have the opposite effect. In commercial practice a satisfactory compromise must be struck between these two factors.

Oxide-coated Cathodes. Experimenters working on the problem of reducing the power required to heat the cathode discovered that, if a pure tungsten filament was coated with an oxide, its emitting properties would be radically changed. Certain oxides, such as that of nitrogen, would greatly diminish its emission, whereas others, such as those of calcium, strontium, and barium, would greatly increase it. Wehnelt developed a filament that, in one of its present forms, consists of a nickel ribbon coated with a mixture of barium and strontium carbonates suspended in a suitable organic binder. In preparing this type of filament the coated ribbon is mounted in the tube, and the bulb is evacuated. The cathode is first heated to perhaps 1300°K for a few moments to reduce the carbonates to oxides and drive off the binder. The surface is then activated by reducing the temperature slightly while a potential applied to the plate draws over the emitted electrons. This plate voltage is maintained until the emission has reached its final value. Although the exact mechanism by which this cathode produces emission is not fully understood, the activation process apparently results in the reduction of some of the oxide to the metallic form from which the emission seems to take place.

When finally formed this type of cathode operates at a temperature of 1100 to 1200°K. This temperature should be carefully maintained as either a higher or lower temperature will shorten the cathode life. Higher temperatures will cause too rapid evapo

ration of the surface while too low a temperature tends to cause concentration of the emission current in small areas, where the temperature will be raised to excessive values and cause deterioration of the surface even though the average temperature of the cathode is subnormal. At normal temperature the emission efficiency is approximately 50 to 150 ma/watt.

Thoriated-tungsten Cathodes. Another type, known as the *thoriated cathode*, was developed by Langmuir[1] and his associates It is made by dissolving a small amount of thorium oxide and a little carbon in a tungsten filament. The filament is first flashed at 2800°K for 1 min to reduce some of the thorium oxide to thorium, then "formed" or activated by being operated *in vacuo* at a temperature somewhat above normal, usually at about 2200°K.[2] At this temperature some thorium works its way to the surface of the filament where it deposits in a layer 1 atom deep. If the process is continued indefinitely, more thorium is brought to the surface, but the layer is never more than 1 atom deep, the additional thorium merely displacing some that has already formed and so gradually reducing the supply of this metal in the cathode. This cathode is ordinarily operated at a temperature of about 1950 to 2000°K Like the oxide-coated cathode, it should not be operated at temperatures either above or below the rated value if maximum life expectancy is to be realized The emission efficiency is approximately 40 to 100 ma/watt.

The emission of the thoriated filament remains normal as long as the layer of thorium is intact; but as the cathode is used, the thorium is gradually dissipated, and after a period of time the emission begins to drop. This decrease in emission is not directly proportional to the decrease in area of the thorium coating but is much more rapid. For example, when the thorium coating has been so reduced as to cover only about one-half of the cathode, the emission will have decreased to less than 1 per cent of its maximum. The minimum emission below which a tube is not usable varies with different tubes but is usually specified by the manufacturer.

Carbonization of Thoriated Cathodes. Both thoriated and oxide-coated cathodes are particularly subject to injury under posi-

[1] I. Langmuir, The Electron Emission from Thoriated Tungsten Filaments, *Phys Rev*, 22, p. 357, 1923.

[2] All other electrodes in the tube should be connected to the cathode or left free during this process.

tive-ion bombardment, especially the latter type. By a process known as *carbonization*, thoriated cathodes can be protected considerably against such bombardment and against too rapid evaporation of thorium at higher cathode temperatures. Before being activated by the procedure described in the preceding section, the cathode is operated at a temperature of somewhat over 1600°K in the vapor of some hydrocarbon such as naphthalene, benzene, or alcohol. The molecules of the vapor decompose upon striking the hot cathode and deposit carbon, which then combines with the tungsten to form tungsten carbide (W_2C). Since tungsten carbide is quite brittle, great care must be taken not to carry the process too far, about 3 per cent carbon being satisfactory.

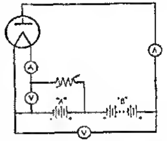


FIG. 1-3. Circuit for determining the emission and space-charge characteristics of a vacuum tube.

The evaporation of thorium from a carbonized cathode is only about 15 per cent as great as that from one not so treated. It is therefore possible to increase the operating temperature of the treated cathode and so increase its emission efficiency. Most high-voltage tubes are supplied with carbonized, thoriated-tungsten cathodes instead of pure tungsten as formerly.

Measurement of Emission. The emission of a tube may be determined experimentally if it is thought to be deficient. A suitable circuit for tubes with thoriated or tungsten filaments is shown in Fig. 1-3. The cathode is heated by a battery as shown or, if preferred, it may be heated by alternating current. A voltmeter is connected across the terminals of the filament, and a rheostat is inserted so that normal voltage may be supplied to the filament. The plate of the tube, together with the grid or grids, if any, is raised to a high positive potential by means of a d-c source marked *B* in the figure. The exact potential of the plate with reference to the negative end of the filament is indicated by a voltmeter as shown. A milliammeter inserted in the plate lead indicates the total plate current; *i.e.*, any electrons emitted by the cathode and passing over to the plate must flow back to the cathode through this milliammeter. The potential of the *B* supply is made sufficiently high that all electrons emitted by the cathode will be drawn over to the plate and none, therefore, will fall back into the cath-

ode. Thus the plate current and the emission current are equal, and the ammeter then indicates the total emission from the cathode.

Tubes having oxide-coated cathodes should not be tested for their emission by the foregoing method, as the very high emission currents that will flow are injurious to the emitting surface. One method that has proved satisfactory with this type of cathode is to apply a short pulse of plate potential and record the emission current photographically by means of an oscillograph, the period of current flow being too short to cause appreciable injury to the emitting surface. For most purposes, however, it is satisfactory to operate the cathode at below normal voltage and observe the emission current by means of an ammeter. Comparison of this current with that of a new tube of the same type, operating at the same reduced voltage, will give a good indication of the degree of normality of the emission of the tube under test. The reduction in cathode voltage should be sufficient to limit the emission current to not much over the normal plate-current rating of the tube.

Another difficulty encountered in testing the emission of oxide-coated cathodes is that the emission is apparently increased by an increase in plate potential. This is due to the irregular surface of the emitting material which produces small pockets from which emission currents will not readily pass except with very high plate potentials This phenomenon may be more readily understood after a discussion of space-charge effects in Chap. 3 (page 43).

Tubes having thoriated filaments that are found to have too low an emission may often be rejuvenated. Low emission is an indication that the surface layer of thorium has been partially dissipated. Generally an additional supply of thorium is still dissolved in the filament in the form of thorium oxide which may be brought to the surface, reduced to thorium, and made available for use by raising the temperature of the filament temporarily. To do this the original activation process is repeated.[1]

A-C Heated Cathodes. Nearly all tubes with rated plate voltage of less than about 1000 are built with the oxide-coated cathode of Wehnelt both for economy of operation and to meet the modern demand for tubes, the cathode of which may be heated directly by

[1] Flashing temperatures are obtained by applying about 3½ times normal filament voltage, and forming or aging temperatures by about 1½ times normal voltage. Flashing is not recommended for tubes having filament voltages higher than 5.

raw (unrectified) alternating current. The use of alternating current creates a demand for a cathode in which the temperature will not respond quickly to changes in filament current. The oxide-coated cathode satisfies this demand more fully than either the tungsten or the thoriated cathodes by virtue of its lower operating temperature and consequent lower rate of heat dissipation.

The oxide-coated cathode is made in two forms: the filamentary type and the indirectly heated type. Filamentary cathodes usually consist of a ribbon of tungsten, nickel, or konal[1] coated with the oxide. Indirectly heated cathodes, which are far more commonly used than the filamentary type, may be constructed as shown in Fig. 1-4, where the cathode consists of an oxide-coated, nickel cylinder, heated by a separate wire or heater carrying the heater current. R is a cylindrical piece of refractory material enclosed by a tubular sleeve of oxide-coated nickel K. This cylinder is heated from within by the heater H, consisting of a tungsten wire operating at a temperature considerably higher than that of the emitting surface K. The temperature of such a cathode is unusually constant and entirely independent of the cyclic variation of the alternating current in the heater wire. Furthermore, all points of the emitting surface are at the same potential, whereas in a filamentary cathode, where the heating current flows directly through the emitting material, there is a difference in potential between any two points in the filament due to the IR drop. Such difference in potential tends to cause an alternating ripple of twice the frequency of the heating current to appear in the plate current of the tube.

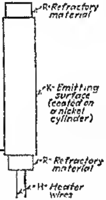


FIG. 1-4. Indirectly heated type of cathode.

Laws Governing Emission. One of the earliest investigations of the laws governing the emission of electrons from hot surfaces was conducted by Richardson in 1901.[2] He developed an equation for electron emission by analogy with the evaporation of atoms from the surface of a liquid. Later a modified form of this equation,

[1] Konal is an alloy of nickel, cobalt, iron, and titanium.

[2] O. W. Richardson, *Proc. Cambridge Phil. Soc.*, 11, p. 286, 1901.

originally suggested by Richardson, was proposed by Dushman,[1] who gives[2]

$$J = AT^2\epsilon^{-\frac{b_0}{T}} \tag{1-2}^*$$

where J = emission current density in amp/sq m of hot surface
T = temperature of hot surface, degrees K
A = constant
$b_0 = \phi e/k$, where ϕ = Richardson's work function, volts; e = electron charge, coulombs, and k = Boltzmann's constant (1.381×10^{-23} joule/deg K)

Table 1-1*

Material	b_0	ϕ, volts	A, amp/sq m
Calcium	26,000	2.24	60.2×10^4
Cesium	21,000	1.81	162×10^4
Nickel	32,100	2.77	26.8×10^4
Oxide coated cathode	11,600	1.00	1.0×10^2
Platinum	59,000	5.08	60.2×10^4
Tantalum	47,200	4.07	60.2×10^4
Thoriated tungsten	33,200	2.86	15.5×10^4
Thorium	38,000	3.35	60.2×10^4
Tungsten	52,400	4.53	60.2×10^4

* Values taken from Saul Dushman, *Elec. Eng.*, 53, p. 1054, July, 1934, and from L. R. Koller, "The Physics of Electron Tubes," McGraw-Hill Book Company, Inc., New York, 1937.

Equation (1-2), commonly known as *Dushman's equation*, is generally conceded today to represent accurately the law of thermionic emission.[3]

Richardson's work function ϕ, which is included in the constant b_0 of Eq. (1-2), is the same factor that appeared in Eq. (1-1). Some common values of ϕ and of A are given in Table 1-1.

For thoriated tungsten, A is a function of the area of the cathode which is covered with thorium, being about 15.5×10^4 for a normal cathode. As the amount of thorium decreases, A tends to increase, approaching the value for tungsten, 60.2×10^4 as the area

[1] Saul Dushman, Electron Emission from Metals as a Function of Temperature, *Phys. Rev.*, 21, June, 1923.

[2] See Preface to the First Edition for meaning of * after equations.

[3] For a more thorough discussion of the factors affecting Eq. (1-2) than that which follows, see Saul Dushman, Electronic Emission, *Elec. Eng.*, 53, p. 1054, July, 1934.

covered by thorium approaches zero. It might be thought from Eq. (1-2) that this would indicate an increase in emission as the amount of thorium decreases, but it must be remembered that b_0 also increases rapidly, and this latter factor is much more effective than A in determining the emission. The emission, therefore, falls off rapidly with loss of the thorium surface. It has also been found that b_0 will increase as the density of thorium covering the surface increases above a certain optimum point. Maximum emission seems to occur when the surface is approximately 75 per cent covered.

While b_0 may be computed from the foregoing information, it is usually determined directly from experimental data. This is done by measuring the emission of an experimental cathode at various temperatures and plotting $\log_e J/T^2$ against $1/T$. The negative slope of the resulting straight line is b_0 and the intercept is $\log_e A$. Values of b_0 are given in Table 1-1 for some of the more common materials.

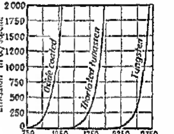


FIG. 1-5. Emission curves for the principal types of emitters as determined by Eq. (1-2).

The curves obtained by applying Eq. (1-2) to each of the three types of emitters previously described are shown in Fig. 1-5. These show that the emission current will increase exponentially without limit as the temperature is increased. In practice the current that may be obtained from a given cathode is limited only by the length of cathode life desired. In Fig. 1-6 the curves of Fig. 1-5 are replotted on power-emission paper[1] with filament power as abscissa and milliamperes of emission as ordinate, the crosses indicating the maximum emission obtainable for each type of emitter with a normal life expectancy.[2] In practice the peak plate current demand should not exceed one-third to one-half of this maximum value.

[1] This paper is sold by Keuffel and Esser, under the name "Power-emission Chart." The use of this special type of cross-section paper results in straight-line emission characteristics. Thus two experimental values will determine a curve, and one or two more will provide adequate checks.

[2] Peak emissions as high as 80 amp/sq cm have been obtained from oxide-coated cathodes, but 0.5 amp/sq cm is a better figure for continuous opera-

Secondary Emission. Suppose an electron to be in an evacuated vessel in which is also a positively charged metal plate. The electron will be attracted to the plate and will "fall" toward it with increasing velocity, striking it with a force depending upon the potential on the plate and the distance through which the electron falls. At the point of impact the kinetic energy of the moving electron will be distributed to other electrons and atoms with which

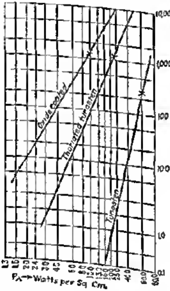


FIG. 1.6. Emission-efficiency curves for the three principal types of emitters plotted on a special power emission chart. The curvilinear arrangement of coordinates is so drawn as to produce straight lines for each emission curve.

the falling electron collides so that these will receive an increase in kinetic energy. If the velocity and, therefore, the kinetic energy of the impacting electron are sufficiently high, one or more other electrons may receive sufficient energy to overcome the potential

---

tion. See John E. Gorham, Electron Tubes in World War II, *Proc. IRE*, 35, p. 295. March, 1947.

barrier of the metal plate and so be emitted from the surface. This phenomenon is known as *secondary emission*.

Secondary emission is not commonly used as a source of electrons for useful purposes, although there are a few applications, such as the electron multiplier (page 151). Nevertheless, emission from this cause exists at the anode of all vacuum tubes. Generally the electrons emitted are attracted back into the anode again by its positive potential and, therefore, have no effect upon the operation of the tube; but if there happens to be a second electrode, close to the source of secondary emission, having a higher positive potential than this source, the secondary-emission electrons will tend to flow to this second electrode. This may readily happen in the four-electrode (tetrode) tube to the detriment of its performance as an amplifier, the flow of secondary electrons *away* from the plate actually being greater, under certain conditions, than the flow of primary electrons *to* the plate.

The effect of secondary emission in these tubes may be nullified by the insertion of an additional, *negative* electrode between the two positive ones. This negative potential will drive all secondary electrons back into the surface from which they were emitted and so obviate the difficulty. It should be noted, however, that the secondary emission is not *eliminated* but is merely prevented from reaching any point where it will interfere with the operation of the tube (see discussion on the pentode, page 78).

High-field Emission. Electrons may be extracted from a conducting surface by the presence of very high electric gradients, of the order of 10^8 volts/m or higher, although such action frequently accompanies some other form of emission, as in the mercury-arc tube. The extremely high gradients required usually exist only in a very thin sheath on the surface of the cathode, a common source of which is the positive ions present in gas-filled tubes (see the first few pages of Chap. 4).

High-intensity fields may also be produced by relatively low potential differences if the electrodes are properly shaped. If, for example, the cathode is constructed of a very fine wire, the immediately surrounding electric field may be very intense. Emission current due to high field intensities is independent of the temperature and is given by

$$J = a\mathcal{E}^2\epsilon^{-m} \tag{1-3}$$

where J = emission current density,
$\mathcal{E}$ = field intensity,
a = a constant
$m = b/\mathcal{E}$, where b is a constant

Photoelectric Emission.[1] Light striking the surface of materials may cause an emission of electrons. This phenomenon is known as *photoelectric emission*. It was discovered by Heinrich Rudolph Hertz in 1887 but could not be put to any commercial purpose for many years because of the small quantity of electrons liberated. With the development of vacuum-tube amplifiers these small currents could be amplified and the photoelectric effect used to control many different operations. Today the photoelectric cell, or "electric eye" as it is often referred to, has revolutionized the motion-picture field, made possible facsimile and television, and provided a means of performing a large number of other processes such as color selection, burglar alarms, and automatic control of lighting.

Hertz's discovery of the photoelectric effect came during the conduct of a set of experiments on spark discharges. The experiments required the operation of two spark gaps, the discharge across one of which was being measured. Hertz noticed that the discharge across the gap under observation occurred more readily when it was illuminated by the discharge across the second gap than when it was not so illuminated. Further experimental work also showed that the effect was produced by the ultraviolet radiation from the second gap and that this radiation must fall upon the negative electrode of the gap under test.

Elster and Geitel and others investigated this phenomenon further and finally established the fact that the cause of the increased discharge across the gap when illuminated was the emission of electrons from the illuminated surface. They also found that with certain of the alkali metals, such as sodium and potassium, electron emission could be obtained by illumination from visible light and even by infrared radiation. This latter discovery laid the foundation for most of the photoelectric applications of today.

Theory of Photoelectric Emission. One of the most striking

[1] An excellent history of the discovery of photoelectric emission and of the development of phototubes (including an extensive bibliography) is given by Alan M. Glover, A Review of the Development of Sensitive Phototubes, *Proc. IRE*, 29, p. 413, August, 1941.

results of the early experimenting was the discovery that the *maximum* velocity of the emitted electrons was entirely independent of the intensity of the impinging light but was a function of its wave length. (The term *light* as used in this discussion is intended to include all radiant energy producing photoemission and not merely visible light.) This was entirely contrary to the classical wave theory of radiant energy, since under that theory the amount of energy absorbed by and, therefore, the velocity of emission of the electron should have been directly proportional to the intensity of the impinging light. Computations indicated that, with light of ordinary intensity, an electron would have to be illuminated for hours before it could absorb a sufficient amount of energy to be emitted against the potential barrier of the material within which it was contained. Experiment, on the other hand, showed that emission started instantly with the illumination of the surface.

As in the case of thermionic emission no explanation for the foregoing phenomenon was found until the development of a new theory. In 1905 Einstein applied Max Planck's quantum theory and postulated that radiant energy existed in small chunks, or photons, rather than in a continuous flow as in the wave theory. These photons contain a definite amount of energy, proportional to the frequency of the radiated wave, being equal to a certain constant h, multiplied by the frequency ν. This constant (known as *Planck's constant*) has been evaluated by a number of experimenters and determined to be $(6.6234 \pm 0.0011) \times 10^{-34}$ joule-sec.[1]

Einstein applied this theory to photoelectric emission by assuming that an electron can receive energy from the impinging light wave only by absorbing a photon. Thus the majority of electrons in an illuminated surface receive no energy whatsoever, but those which are in the path of an impinging photon receive sufficient energy to be instantly emitted, provided the energy contained in the photon is equal to or exceeds that which the electron loses in passing through the potential barrier of the emitting material. If the electron has zero velocity at the time it absorbs a photon and encounters no collisions before being emitted, its energy content after emission should be that of the photon ($h\nu$) less the surface work (ϕe) of the emitting surface, or

$$\tfrac{1}{2}mv^2 = h\nu - \phi e \qquad (1\text{-}4)^*$$

[1] *Revs. Modern Phys.*, 20, p. 106, 1948.

where m = mass of an electron = 9.1×10^{-31} kg
v = velocity of emission of the electron, m/sec
h = Planck's constant = 6.62×10^{-34} joule-sec
ν = frequency of the impinging light, cycles/sec
ϕ = work function of emitting surface, volts
e = charge on an electron = 1.60×10^{-19} coulomb

Threshold Frequency. Referring to Einstein's equation, it is evident that the velocity of emission is just zero when $h\nu = \phi e$; *i.e.*, when the energy of one photon is just equal to the energy lost by the electron in being emitted. This latter energy is a function of the emitting material, since experiment has shown that for every kind of emitting material there is a definite value of ϕ which is the same as that for thermal emission.

Evidently, if ϕ has a definite value, there must be some frequency of light below which no emission will take place from any given material. This frequency is known as the *threshold frequency* and is given by Eq. (1-4) when $v = 0$. Millikan performed the first conclusive experiments to determine its existence and so provide some verification of Einstein's equation. By a series of tests he demonstrated conclusively that, for any given photoelectric material, there is a definite frequency of light below which no emission will take place no matter how great the intensity of the light or how long it may be applied.

Photoelectric Emission from Composite Surfaces. A more sensitive surface and one that will respond to a longer wave length may be obtained by preparing on a metal base a surface layer composed largely of the oxide of an alkali metal plus some of the free alkali metal. One of the most common emitters of this type is cesiated silver, formed by shaping a silver (or silver-coated) surface to the desired form and then cleaning it thoroughly both by solvents and by heat-treatment in a vacuum. The silver is then oxidized by admitting a small amount of oxygen and initiating a glow discharge between the silver surface and another electrode. After the oxidation has proceeded to the desired point, cesium is admitted, and the tube is then baked at a temperature of about 200°C. The resulting surface apparently consists of cesium oxide interspersed with silver and cesium, with a monomolecular layer of cesium over the entire surface. It is often referred to symbolically as Cs-CsO-Ag.

Effect of Light Intensity. As was pointed out in a preceding

section, an increase in light intensity causes an increase in emission but not in maximum velocity of emission. According to the quantum theory an increase in light intensity means an increase in the number of photons and, therefore, in the number of electrons that can absorb energy and be emitted. This relationship is a direct proportionality so that the relative intensity of a light beam may be determined by observing the emission current produced from a given surface. The principle is made use of in many modern applications such as photographic meters, illumination meters, sound-on-film talking pictures, and facsimile.

Spectral Selectivity. It might seem from Einstein's equation that all materials would be sensitive to any light having a frequency higher than the threshold value. As a matter of fact this is not the case, the curve of emission vs. wave length for most materials having one or more peaks (Fig. 1-7).[1] This phenomenon is known as *spectral selectivity*. Actually the response of most emitters is also affected by polarization of the light, being more responsive to one direction of polarization than the other. This phenomenon is known as *polarization selectivity*. In most commercial tubes the emitting surface is too rough to permit any differentiation between this phenomenon and that of spectral selectivity.

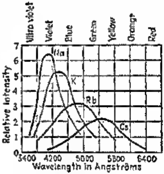


FIG. 1-7. Curves illustrating the spectral selectivity of various photoelectric materials.

Most of the energy in the usual sources of artificial light lies toward the red end of the spectrum. Inspection of Fig. 1-7 leads to the conclusion that cesium is the most satisfactory material for phototubes that are to be used with such sources of light, since its maximum response is more nearly toward the red than that of the other elements. For some time cesium cells were used almost exclusively in commercial applications.

The development of composite surfaces (*e.g.*, the cesiated-silver surface described in a preceding section) has produced phototubes with response running well into the infrared. Such cells com-

[1] 1 angstrom $= 10^{-10}$ meter $= 10^{-4}$ micron.

monly have a threshold wave length as long as 12,000 or 13,000 A, and special emitters have been constructed responsive to 17,000 A. Composite surfaces usually produce two peaks in the responsive curve, the second one occurring in the ultraviolet region.[1]

Problems

1-1. Compute the minimum velocity that an electron must attain in order to be emitted from (a) thorium, (b) tantalum, (c) calcium.

1-2.* Using Dushman's equation, calculate and plot a curve of emission current vs. temperature for a tungsten filament 10 mils in diameter and 3 in. long.

1-3.* Repeat Prob. 1-2 for a thoriated-tungsten filament.

1-4.* Repeat Prob. 1-2 for an oxide-coated filament.

1-5. A two-element vacuum tube is operated with 50 volts between cathode and anode. How much energy is imparted to the anode by each electron flowing from cathode to anode? (Assume that the electron has zero velocity after emission from the cathode.)

1-6. Compute the threshold frequency of light for (a) calcium, (b) cesium.

---

[1] Additional spectral response curves will be found in Chap. 5.

* A high degree of accuracy is necessary in determining $\epsilon^{-b_0/T}$ in these problems. The use of logarithms rather than slide rules is recommended.

CHAPTER 2

CONSTRUCTION OF VACUUM TUBES

Details of construction of vacuum tubes vary somewhat with the manufacturer. However, general principles and materials of construction are much the same, irrespective of the make of the tube. The purpose of this chapter is to provide the reader with an idea of the kind of materials going into the construction of vacuum tubes and with typical methods of assembling and evacuating them.

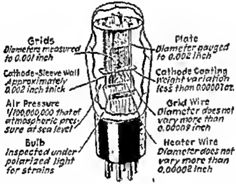


FIG. 2-1. Cross-section view of a glass-envelope tube showing the various parts. (*RCA*.)

General.[1] Basically, vacuum tubes consist of a cathode and an anode, with usually one or more grids, all mounted on a suitable supporting structure and contained in a vessel of glass or metal. This vessel is evacuated to a high degree, although in certain tubes inert gas or mercury vapor is then inserted to a low pressure. Tubes in which gas has been inserted are known as *gas-filled tubes*; those which contain no gas are known as *high-vacuum tubes*.

Figure 2-1 is a cross section of a glass-envelope, high-vacuum, triple-grid tube of low power (such as is used in radio receivers or

[1] See Appendix A for definitions of unfamiliar words.

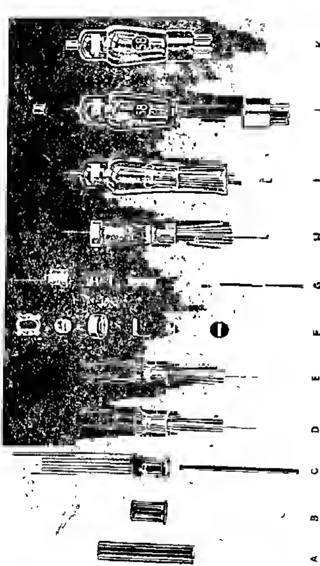


Fig. 2-2 Structure of a glass-envelope tube (RCA)

voltage amplifiers). The cathode is of the indirectly heated type although largely obscured by the grid wires surrounding it. All the electrodes are supported by wires held in a glass "pinch" at the base of the tube and by a mica disk at the top which fits snugly into the dome at the top of the glass envelope.

The method of assembling such a tube is illustrated in Fig. 2-2.[1] The glass tube shown at A is worked down in a lathe by applying gas flames to form the pinch shown at C. The supporting wires are then inserted, and the glass is pressed together, while hot, to form a gastight seal, as at D. The ends of the supporting wires inside the envelope are made of nickel; the ends that are sealed through the press are made of copper-coated nickel-steel, known

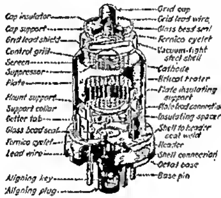


FIG. 2-3. Cutaway drawing of the 6J7, all-metal, screen-grid pentode, showing component parts. (*RCA*.)

as *Dumet*, which has about the same coefficient of expansion as the glass used.

The electrodes are illustrated at G. From top to bottom these parts are anode, suppressor grid, screen grid, control grid,[2] cathode, and heater. View F shows some of the small parts used in the tube. From top to bottom these are: mica disk which fits in the dome of the glass envelope to give increased rigidity to the tube, mica spacer and insulator which fit over the top of the supporting wires, shield for the grid wire which runs to the cap on the top of

[1] This is a triple-grid tube of a type slightly different from that of Fig. 2-1, but the essential elements are the same.

[2] For the purpose of these various grids and the anode, see Chap. 3.

the tube, cup to hold the getter, a second mica spacer, and the grid collar.

View *H* shows the completed assembly, and *I* shows the assembly inserted into the glass envelope. The bottom of the pinch is sealed to the envelope, and the tube is evacuated. It is then ready for the base and grid cap as at *J*, with the finished tube shown at *K*.

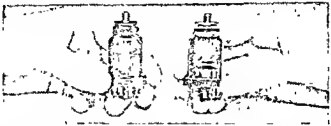


FIG. 2-4. A cutaway view of a metal-envelope, pentode tube. The supports for the three grids and the upper end of the cathode may be plainly seen above the plate. (*General Electric Co.*)

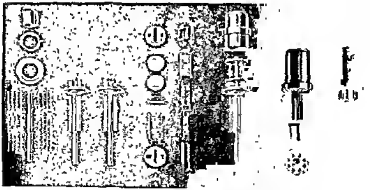


FIG. 2-5. Structure of a metal envelope tube. (*RCA.*)

Many tubes are now made with metal envelopes, as shown in Figs. 2-3 and 2-4. Here the supporting wires are inserted in individual glass beads located around the periphery of the metal base. Thus these wires are more widely spaced than in the glass-envelope tubes, reducing the internal capacitances. The method of assembling such a tube, together with details of its construction, is shown in Fig. 2-5.

Two tubes of the glass-envelope type are shown in Figs. 2-6 and 2-7, and a metal-envelope tube is shown in Fig. 2-8.

Cathodes. Methods of forming the composite cathodes (oxide-coated and thoriated tungsten) were described in Chap. 1. As

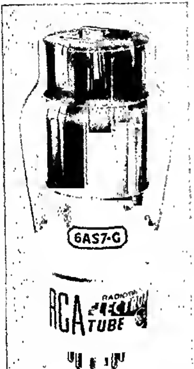


FIG. 2-6 Glass-envelope tube (low-mu, twin power triode). (*RCA.*)

FIG. 2-7. Glass-envelope tube (twin-triode amplifier). (*RCA.*)

there pointed out, each type of cathode has its own field of usefulness. Thus the oxide-coated cathode has the highest emission efficiency (ratio of emission current to power required to heat the cathode, usually measured in milliamperes per watt) and is therefore used wherever it is feasible to do so. Unfortunately it is more vulnerable than the other common types and may not be used

where there is danger of appreciable positive-ion bombardment. Therefore, it is used only in tubes with a comparatively low tube drop. Its use with high-vacuum tubes is confined to those with rated anode voltages not much in excess of 600. Theoretically there should be no gas in a high-vacuum tube and thus no positive-ion bombardment, but practically it is impossible to evacuate any vessel completely; ionization of the remaining gas may then take place at high plate potentials with consequent damage to an oxide-coated cathode. Oxide-coated cathodes are used in virtually all gas-filled tubes, since the tube drop is kept so low (by the presence of the ionized gas) that positive ions cannot attain high velocities sufficient to damage the cathodes.[1] Cathodes of gas-filled tubes are commonly heat-shielded to reduce the heating power required.[2]

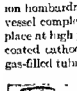


FIG. 2-8.—Metal-envelope triode (*RCA*).

The thoriated cathode is used for tubes of higher voltage rating than is safe for oxide-coated cathodes. It is less subject to damage than the oxide-coated cathode and will withstand considerable positive-ion bombardment, especially since the development of the carbonization process (see page 10).

Tungsten filaments have long been used in all high-vacuum tubes with very high plate voltages. However the thoriated-tungsten filament has now been developed to the point where it has almost entirely replaced the pure tungsten.

Great care must be used in assembling a vacuum tube to avoid contamination of the cathode surface, especially when composite cathodes are being used. Workers assembling them, for example, must use care not to touch the cathodes with their fingers, as the natural skin oils will greatly affect performance.

The smaller tubes, such as are used in radio receivers, public-address systems, and various types of industrial control equipment, are now made almost exclusively with indirectly heated cathodes, while filamentary cathodes are used in many of the larger tubes.

Grids. The grids of high-vacuum tubes are usually made of fine wire wound laterally in grooves on the supporting wires. A

[1] See Chap. 4 for an explanation of this phenomenon.

[2] See p. 104 for a discussion of heat-shielded cathodes.

swaging process is then used to hold the grid wires in place. Manganese nickel is commonly used for the grid wires of receiving-type tubes, while zirconium- or platinum-clad molybdenum or gold-plated molybdenum wires are used for transmitting tubes.[1] With these materials primary grid emission is very low, a factor of great importance in most high-power tube applications.

The grids of gas-filled tubes commonly consist of a metallic cylinder containing a baffle plate mounted normal to the axis of the cylinder. This baffle plate is perforated to permit passage of the electrons (see pages 120 and 121).

Anodes. The anodes of small, high-vacuum tubes are made of nickel or iron. They are pressed out of sheet material which is frequently crimped or flanged to increase rigidity. The larger sizes are blackened to increase the radiation of heat. The anodes of power tubes (those with plate voltages in excess of perhaps 500 volts) are often made of graphite because of its superior performance under high temperature conditions. Zirconium or zirconium compounds are sometimes sprayed onto metal anodes to improve their black-body radiation properties and to serve as a getter (see Getter, page 34).

The anode in a gas-filled tube is usually a small carbonized nickel button located above the cathode, often partly enclosed by the cylindrical grid structure (in those tubes containing a grid) as shown in Fig. 4-22 (page 120). The area of the plate need not be so large as is required for a high-vacuum tube, since the tube drop and, therefore, the heat dissipation at the plate is much less (see discussion on page 101).

Figure 2-9 shows two gas-filled triode tubes, known as *thyratrons*, *a* being a glass-envelope tube and *b* one with a metal envelope. Cooling fins may be seen attached to one end of the metal-envelope tube to increase the heat-radiating surface.

Figure 2-10 shows a number of glass-envelope, gas-filled diodes of various plate-voltage ratings.

Methods of Cooling. Cooling of most tubes is accomplished by natural circulation of air around the envelope. This means that the heat liberated at the anode must pass through the vacuum inside the tube and through the glass or metal envelope before being

[1] George A. Espersen, Fine Wires in the Electron-tube Industry, *Proc. IRE*, 36, p. 116W, March, 1948; John E. Gorham, Electron Tubes in World War II, *Proc. IRE*, 35, p. 295, March, 1947.

carried away by the cooling air. For tubes with ratings up to 1 kw this offers no serious problem, but in the larger sizes of high-vacuum tubes it is impossible to dissipate the heat in this manner without an excessive temperature rise, some of the earlier high-power tubes in glass envelopes having become so hot that the glass collapsed under atmospheric pressure. Thus tubes with ratings in

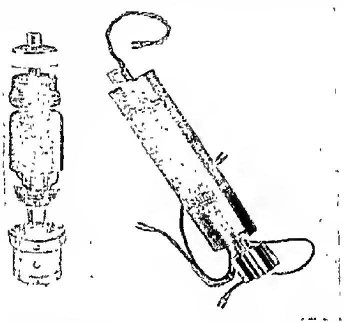


FIG. 2-9. Two grid-controlled, mercury-vapor tubes (thyratrons). Rating of tube a: average plate current = 1.6 amp., peak plate current = 6 amp., inverse peak voltage = 10,000. Rating of tube b: average plate current = 12.5 amp., peak plate current = 100 amp., inverse peak voltage = 2000. (Westinghouse Electric Corp.)

excess of about 1 kw require special means of cooling, one of which is immersion in a water jacket. The tube of Fig. 2-11 is of this type, the metal cylinder that constitutes the plate also serving as a part of the envelope. The glass structure at the top supports the grid and filament which extend down into the anode. The

seal between the glass and metal portions of the envelope is made by using materials that have nearly the same coefficient of expansion and by feathering the edges of the anode. Water cooling is provided by inserting the anode into a water jacket (Fig. 2-12) through which water is circulated. The water is cooled in a suitable manner, as by passing it through a radiator across which air is blown.

Another method is to provide the metal anode with radiating fins, as in Fig. 2-13, and blow air across the fins. Installation costs

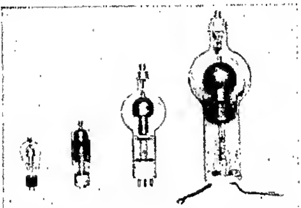


FIG. 2-10. Mercury-vapor tubes of various sizes. (*General Electric Co.*)

with this type of tube tend to be less than with water cooling but, because of the heavy fin structure, packing and shipping of air-cooled tubes are more difficult and costly than for water-cooled tubes. Air-cooled tubes are particularly advantageous in installations where no heat is supplied to the building and ambient temperatures may drop below freezing or where there is no adequate supply of pure fresh water, as on shipboard.[1]

Demountable tubes have been used to some extent in Europe, especially in France. Tubes of this type use a water-cooled plate but can be disassembled and repaired whenever the cathode burns out. This procedure avoids the problems of making airtight glass-to-metal seals but requires continuous evacuation. The principal

[1] For a discussion of the relative merits of water and air cooling, see I. E. Mouromtseff, Water and Forced-air Cooling of Vacuum Tubes, *Proc. IRE*, 30, p. 190, April, 1942.

disadvantages of this type of construction are the time required to put equipment back into service after a tube failure and the need for personnel of sufficient skill to maintain the pumping equipment. Thus far, sealed-off tubes have proved more satisfactory in sizes up to about 500 kw, but it appears that even larger tubes are

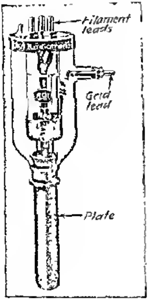


FIG. 2-11. Water cooled high-vacuum tube (RCA.)

FIG. 2-12. Water jacket into which the tube of Fig. 2-11 is inserted (RCA.)

needed, especially for industrial use, and these will quite likely be of the demountable type.[1]

The energy lost at the anode of gas-filled tubes is sufficiently small that water cooling is not normally required. If the tube is to be placed in a closed space, forced-air circulation may be neces-

[1] I. E. Mouromtseff, Development of Electronic Tubes, *Proc. IRE*, **33**, p. 231, April, 1945; I. E. Mouromtseff, H. I. Dailey, L. C. Werner, Review of Demountable vs. Sealed-Off Power Tubes, *Proc. IRE*, **32**, pp. 653–664, November, 1944.

sary; otherwise natural air movements will usually provide adequate cooling, except in the largest sizes.

Evacuation of Vacuum Tubes. All tubes, even those which are to contain gas, must be highly evacuated. The presence of any gas in what is intended to be a high-vacuum tube results in erratic behavior due to ionization of the gas under the impacts of the

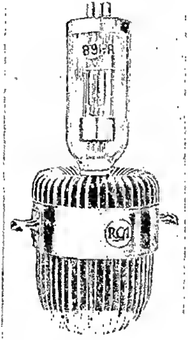


FIG. 2·13. Tube designed for forced-air cooling. (*RCA.*)

emitted electrons. Those tubes which are intended to contain gas must first be highly evacuated in order that they may contain only the desired gas, oxygen and water vapor being especially objectionable in tubes.

Excellent pumps are now available for performing this evacuating process. Mere pumping is, however, not sufficient, as the metal parts and glass parts are capable of absorbing a large amount

of gas This is driven off by operating the tube at a high temperature during the evacuating process. Tubes with glass envelopes are generally heated by inserting them in a r-f field which sets up eddy currents in the metallic structure.[1] Metal-envelope tubes are heated by gas flame or other similar means, since the r-f method would be no more effective in heating the internal parts than is the gas flame, owing to the short-circuiting action of the envelope. The tube must be on the pump during this heating process, as it may require several hours to evacuate a large tube.

A complete vacuum is, of course, impossible of attainment. Even with the highest vacuum attainable there are billions of atoms within the tube Nevertheless the atoms are so small that an electron may travel many times as far as the distance between the cathode and plate without colliding with one, and as long as collisions are very infrequent, the gas remaining in the tube is not injurious The probability of an electron striking an atom in its travel across the tube is measured in terms of the *mean free path* of the electron (see also page 19), *i.e.*, the average distance that it may travel without collision As long as the mean free path is much greater than the distances between electrodes, the vacuum is satisfactory for commercial operation

Getter. Even with the highest refinement of the evacuating process possible today, gas may be liberated in the tube while in service Sudden overloads or even long service at normal loads may permit some of the occluded gas that was not removed during the manufacturing process to escape into the bulb. Consequently all tubes contain material, known as a *getter*, that is capable of readily absorbing gas Examples are cerium, aluminum, magnesium, barium, and red phosphorus.

It has been found that the most satisfactory getter is one made of the same material as the emitting surface Therefore, barium is widely used in this capacity today, since a large majority of tubes are made with oxide-coated cathodes Unfortunately barium, unlike magnesium for example, is not protected by its own oxide, and new methods had to be developed for its use. Most of them consist either of supplying the barium in the form of a compound that will break down or otherwise produce barium when heated or of enclosing the metallic barium in a metal tube that is

[1] Edwin E Spitzer, Induction Heating in Radio Electron-tube Manufacture, *Proc IRE*, 34, p 110W, March, 1946.

pinched nearly airtight at the ends.[1] All such methods require flashing of the getter in the tube by application of heat such as may be supplied by focusing a r-f field on the getter through the glass envelope and are, therefore, not suitable for use with metal-envelope tubes. A suitable method for metal-envelope tubes uses a tantalum ribbon formed into the shape of a trough and filled with barium berylliate.[2] This ribbon is then welded between the ground pin of the octal base and the shell. A current sent through the tantalum between the shell and the pin will cause the barium berylliate to react with the tantalum, producing free barium.

Fine zirconium wire has been used as a getter in X-ray tubes where the higher vapor pressure of barium makes that material less satisfactory. The zirconium wire is wound alongside a slightly larger tungsten wire on a tungsten or molybdenum core and is then connected into the filament-heating circuit so that its temperature is maintained at from 1300 to 1900°K. In this range of temperatures it will absorb all gases except the inert gases.[3]

Care must be taken that the getter does not condense on parts of the tube where it will reduce the resistance between terminals, *e.g.*, at the top of the bulb of a screen-grid tube, if the grid lead is brought out at that point. Also, the glass envelopes of high-power tubes must not be covered any more than necessary, as the mirror-like surface of the condensed material will reflect a large part of the radiant energy back into the tube and so prevent satisfactory cooling under load. The getter is, therefore, so located that, when flashed, the barium (or other material) will deposit on a small portion of the envelope where it will not cause trouble. The trough type of construction, described in the preceding paragraph (see the fifth component from the top in the fourth column from the left in Fig. 2-5), or the use of a small cup in glass-envelope tubes (see fourth component from the top in column *F*, Fig. 2-2) makes it possible to direct the evaporation of the getter in any desired direction and over a very small area. This is in marked contrast to the early thoriated-filament tubes (*e.g.*, '01A) where the getter covered almost the entire envelope.

[1] E. A. Lederer and D. H. Wamsley, "Batalum," A Barium Getter for Metal Tubes, *RCA Rev.*, **2**, p. 117, July, 1937.

[2] E. A. Lederer, Recent Advances in Barium Getter Technique, *RCA Rev.*, **4**, p. 310, January, 1940.

[3] Espersen, *loc. cit.*

CHAPTER 3

HIGH-VACUUM TUBES

Emission of electrons from a cathode contained in an evacuated vessel does not of itself constitute a useful phenomenon; but if an additional electrode (or anode) is inserted into the vessel, raised to a potential somewhat more positive than the cathode, electrons will be attracted to this electrode and so constitute a current flow through the evacuated space.[1] The number of electrons passing over to the anode is determined either by the emission of the cathode or by the potentials existing in the region between the cathode and anode or by both. The characteristics of this electron flow in various types of tubes are presented in this and the next three chapters.

1. DIODES

Tubes containing a cathode and one additional electrode (known as the *anode*, or *plate*) are called *diodes* or *two-element tubes*. The anode of a high-vacuum tube ordinarily consists of a conducting surface which completely surrounds the cathode.[2] It does not normally emit electrons but serves only to attract them from the cathode.[3]

[1] Note that electrons flow from cathode to anode but that our generally assumed positive direction of current flow is from anode to cathode (see p 7, Chap 1)

[2] See Chap. 2 for further details of tube construction.

[3] An anode may emit electrons under certain conditions. The most common cause is that of secondary emission which exists in most tubes Since this type of emission takes place only when the plate is being bombarded by electrons and therefore when it is positive with respect to the cathode, all electrons so emitted will tend to fall back into the anode and thus produce no noticeable effect, unless another more positive electrode is present as in tetrodes (pp 70–77) Sometimes cathode material is distilled over onto the plate either during manufacture or while the tube is in service. If the plate then becomes very hot, as it may under heavy loads,

Probably the earliest diode was the special lamp with which Edison discovered what has since been termed the *Edison effect* (page 5). Later, in 1905, Fleming patented a similar device known as the *Fleming valve*, to be used as a detector of radio currents. The modern diode differs little from these early models in so far as its principle of operation is concerned but is much improved in increased ruggedness, longer life, and higher vacuum.

Characteristic Curves of High-vacuum Diodes. As was seen in the preceding chapter, the number of electrons emitted by a cathode is dependent upon the temperature and therefore upon the current flowing in the heater. This can be shown experimentally by means of the circuit of Fig. 3-1.[1] It is best to use a tube with a filamentary, rather than an indirectly heated, cathode, since the cathode temperature of the latter reaches a steady value only after a considerable period of time. The plate potential should be kept constant at any desired value, the filament voltage varied in steps, and the plate current noted. The process may be repeated for any number of different plate voltages. A set of curves taken in this manner is shown in Fig. 3-2. By cross reading these curves another set may be plotted with e_b as abscissa[2] as in Fig. 3-3.

The *exact performance of any tube* under any given impressed emfs *may be predicted from a set of these curves*, provided the potential between plate and cathode does not change appreciably during the time an electron is crossing the space between these two electrodes. This limitation is usually expressed by stating that the *electron transit time* must be short compared to the time of one cycle of the alternating voltage impressed on the tube. For all practical purposes this means that at frequencies less than a few hundred megacycles per second (or even somewhat higher with

some emission may take place by thermionic action. Also some anodes have been known to exhibit photoemissive properties.

Generally speaking, emission from anodes is to be avoided as far as possible and particularly if it tends to occur while the plate is more negative than the cathode (or more negative than any other electrode in tubes containing additional electrodes).

[1] Batteries are frequently used in this book to indicate a source of direct potential. In most cases the source of potential actually used in practice would be a vacuum-tube rectifier (Chap. 7) or, occasionally, a d-c generator.

[2] The reader should refer to Appendix A for interpretation of the notation used.

suitably designed tubes) the instantaneous plate current may be read directly from the curves whenever the instantaneous plate voltage and steady value of filament voltage (or cathode heater voltage) are known

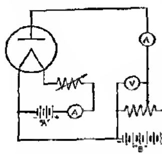


FIG 3 1 Circuit for determining the characteristic curves of a diode

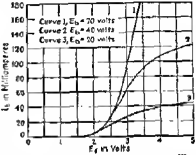


FIG 3 2 Static characteristic curves of a diode, i_b vs E_f

The curves of Fig 3-2 may also be obtained with 60-cycle alternating current flowing in the cathode heating circuit. With filamentary-type cathodes the resulting curves will be essentially the same as those shown if the plate voltage is decreased by approximately half the filament voltage. Since the filament voltage, under a-c operation, reverses polarity twice in each cycle, the effect is much the same as if the negative terminal of the plate battery in Fig. 3-1 were connected alternately first to the negative terminal and then to the positive terminal of the filament and an average value of plate current was obtained. This in turn is essentially equivalent to connecting the plate return to the mid-point of the filament. If the plate return lead were so connected, it is obvious that the average potential difference between plate and cathode would be increased by half the filament voltage. Decreasing the plate supply voltage by this amount will evidently reestablish the original conditions.

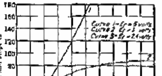


FIG 3-3 Static characteristic curve of a diode, i_b vs e_b

Where indirectly heated cathodes are used, there will obviously be no difference between a-c and d-c operation of the cathode heating circuit.

Saturation and Space Charge. The plate current flowing in a diode under any given conditions is always determined by one or both of two limiting factors. One of these factors is imposed by the maximum emission of the cathode and is termed *saturation*. The other is imposed by the accumulation of negative electrons in the space around the cathode and is known as *space charge*.

To obtain a simple picture of these two factors, let a cathode and plate be assumed in the form of two infinite parallel planes. If the cathode is cold and zero potential is applied between these two planes, the curve of potential distribution in the intervening space is a straight horizontal line, as shown in curve a, Fig. 3-4, where K represents the cathode, P the plate, and KP the distance between cathode and plate.

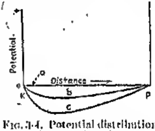


FIG. 3-4. Potential distribution curves for $e_b = 0$.

Next, suppose the cathode to be heated to its normal temperature. The plate, having no positive charge, does not attract the emitted electrons; therefore they fall back into the cathode as rapidly as they are emitted.[1] The cloud of electrons thus formed around the cathode constitutes a negative charge, known as the *space charge*. The potential distribution curve is therefore not a straight horizontal line but is a curved one as shown in curve b, Fig. 3-4, the amount of negative dip being determined by the space charge.

With zero plate potential, the space charge will always be such as to just counteract the emission. Thus, if the emission is increased by raising the cathode temperature, the number of emit-

[1] Actually a few will be emitted with sufficient velocity to be carried over to and strike the plate, whence they will return through the external circuit. The number of electrons thus passing to the plate will be very small, since only those having the very highest velocity will be able to penetrate the electron cloud to such an extent. Had the plate been left free instead of being maintained at zero potential, electrons would have accumulated thereon until the plate became so negative that no more could overcome the increased negative gradient, and the electron flow would have ceased entirely.

ting electrons and the magnitude of the space charge both increase until equilibrium exists, the electronic flow to the plate still being essentially zero. This is shown by curve c, Fig. 3-4, which represents the potential distribution with a higher cathode temperature.

Next let it be assumed that the potential of the plate is raised to the positive value xy, Fig. 3-5, with normal cathode temperature. If the cathode were cold, the potential distribution within the tube would again be a straight line (curve a); but since the cathode temperature is normal, a cloud of electrons exists in the space surrounding the cathode, again producing a negative dip in the curve. However, the plate now exerts a force of attraction on the emitted electrons, and a large number of them are drawn over to it. This reduces the number of electrons in the space between the two electrodes, or, in other words, the space charge is partially neutralized by the positive plate potential. The resulting potential distribution curve is somewhat as shown in curve c, Fig. 3-5. The negative slope of the curve close to the cathode shows that electrons emitted with low velocities are unable to overcome the negative charge mn and must return to the cathode, while those with somewhat higher emissive velocities are able to pass over the negative "hump" and so "fall" into the plate. The maximum negative charge mn in this case is of necessity less than the corresponding maximum of curve b, Fig. 3-4, since that of Fig. 3-4 was sufficient to prevent virtually all electrons from passing over to the plate.

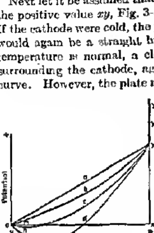


FIG. 3-5. Potential distribution curves for $e_b = xy$

If the cathode is heated to a higher temperature, the emission will be increased, thereby increasing the space charge and the negative gradient at the cathode but without appreciably increasing the current to the plate. This condition is shown by d, Fig. 3-5.

If the cathode temperature is reduced, the electron emission is decreased until finally a condition similar to b, Fig. 3-5, is reached where the potential gradient from cathode to plate is positive all

the way, so that all electrons emitted will pass over to the plate. The plate current flowing under this condition will necessarily be less than for curves *c* and *d*.

General Shape of Static Characteristic Curves. On the basis of the curves of Fig. 3-5 it is now possible to predict the shape of the curves of Fig. 3-2. The results of such a prediction are shown in Fig. 3-6, for a theoretical tube with infinite, parallel-plane electrodes, and the method by which they were obtained follows. The reason for the slight difference between these curves and those of Fig. 3-2 will be explained.

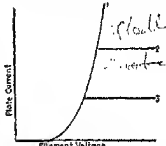


FIG. 3-6. Theoretical static characteristic curves for infinite, parallel-plane electrodes. i_b vs. E_f.

If it is assumed that a high plate potential is applied to a high-vacuum diode with zero filament potential, there will necessarily be no plate current, since there will be no emission. If the filament potential is then gradually increased, a point will finally be reached where emission will begin and a small plate current will flow. With further increase in E_f, the emission will increase exponentially in accord with Eq. (1-2); and, if the plate potential is sufficiently high to attract all emitted electrons, the plate current will be always equal to the emission current. We can, therefore, draw curve 1, Fig. 3-6, to represent this condition. At all points of this curve the plate current is equal to the emission or saturation current so that it is determined entirely by the first of the two limiting factors, saturation. The potential distribution curve for this condition will be a straight line (*a*, Fig. 3-5) when the cathode is cold and will become more and more curved as the cathode temperature is increased. Nevertheless, since the plate potential is at all times sufficiently high to draw over all emitted electrons, the potential gradient must be positive at all points within the tube. The potential curve, therefore, never crosses the zero line, as does *c* of Fig. 3-5, but must be similar to *b* of the same figure.

With a low plate potential, the plate current will still increase according to Eq. (1-2) as E_f is increased up to a point at which the

plate potential can no longer draw over all the emitted electrons, i.e., cannot completely neutralize the negative space charge at the cathode. The potential distribution curve at this point will be just tangent to the zero line at the cathode (Fig. 3-5). Any further increase in E_f will cause the potential to be negative for a portion of the space immediately adjoining the cathode, as in c, Fig. 3-5, and the curve of i_b vs. E_f will flatten out as in curve 2, Fig. 3-6, with the limiting factor being *space charge*.

Fig. 3-7. Small double anode diode.

Curve 3, Fig. 3-6, represents a still lower value of e_b, showing similar results but with a lower maximum value of i_b.

Summarizing, the current along the flat portions of the curves of Fig. 3-6 is limited by space charge and along the curved portions by saturation.

A comparison of the experimental curves of Fig. 3-2 with the theoretical ones of Fig. 3-6 will show a strong similarity, but the experimental curves are much less abrupt. The spacing between the plates and cathodes of commercial tubes is not the same at all points as in the case of the theoretical infinite planes considered, so that at certain values of E_f some portions of the electron flow within the tube may be limited by space charge and other portions by saturation, giving a much less abrupt change. The tube used to obtain the curves of Fig. 3-2 was an 80 (see Fig. 3-7) having two V-shaped filaments, each surrounded by a plate in the shape of a rectangle. Both filaments and both plates were connected in parallel, thus serving as a single tube.

If the same line of reasoning is applied to the curves of i_b vs. e_b, the theoretical curves of Fig. 3-8 may be drawn and compared with those of Fig. 3-3. It should be noted that, although these curves are very similar to those of Fig. 3-6, the current is limited by *saturation* for the flat portions and by *space charge* for the curved portions, just opposite to the conditions of Fig. 3-6. The reader should check this point thoroughly by analyzing these curves in the same manner as those of Fig. 3-6 were analyzed.

Another observable difference between theoretical and experimental curves may be seen in comparing Figs. 3-3 and 3-8 in that curve 3 of Fig. 3-3 never does become quite flat. This is characteristic of oxide-coated cathodes (such as used in the 80 tube on which these curves were obtained). This effect is caused by irregularities in the surface of the oxide. If this surface could be magnified sufficiently, it would be found to contain many pockets, or "canyons." The pockets are of a higher temperature than the outer surface, because of their lesser exposure, and consequently will saturate only with a higher plate voltage. Furthermore, electrons accumulate in these pockets and produce a strong space charge which the plate can overcome only at very high potentials. Pure tungsten filaments, on the other hand, have a much smoother surface, and curves taken on a tungsten-filament tube show a considerably flatter saturation curve as well as a more abrupt curvature as saturation is approached.

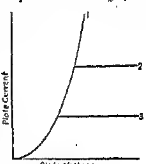


FIG. 3-8. Theoretical static characteristic curves for infinite, parallel-plane electrodes. i_b vs. e_b.

A mechanical analogy of the effect of space charge is shown in Fig. 3-9 which represents a surface shaped like the potential distribution curve c of Fig. 3-5, inverted. If balls, representing electrons, are rolled along the line HK and "emitted" at K at varying velocities, some will have sufficient energy to pass over the space-charge "hump" in the surface and so fall into the plate, whereas the others will fall back onto the plane representing the cathode. If the balls are increased in number and average velocity and the height of the hump is increased a corresponding amount, the number of balls passing over the hump will remain the same, giving the conditions of curve d, Fig. 3-5. But if the hump is eliminated entirely, all the balls will go over to P, showing the effect of saturation, as in curve b, Fig. 3-5. The reason that the curve must be inverted is that electrons bear a negative charge and are, therefore, attracted, or "fall," toward positive charges.

Space-charge Equation. The equation of the plate current in

a diode when limited by space charge only (such as curve 1, Fig. 3-3) has been shown by Child[1] to be

$$I_b = \frac{KE_b^{\frac{3}{2}}}{x^2} \tag{3-1}^*$$

where I_b = space or plate current
E_b = plate voltage
x = distance between electrodes
K = constant, depending upon geometry of tube

Unless one is interested in vacuum tubes from the standpoint of the designer, Eq. (3-1) will probably be of more interest if the

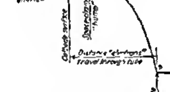


FIG. 3-9. Mechanical analogy of electron emission. Balls rolling past point K represent emitted electrons.

spacing factor x is absorbed in the constant, since the spacing between electrodes is not adjustable except in manufacture. With this simplification Eq. (3-1) becomes

$$I_b = K_1E_b^{\frac{3}{2}} \tag{3-2}^*$$

where $K_1 = K/x^2$.

The plate current in a diode is seen from Eq. (3-2) to vary as the three-halves power of the plate voltage. Actually this theoretical relationship of Child's depends on certain assumptions: (1) that the electrodes are parallel infinite planes, (2) that the plate current is limited only by space charge (not by emission saturation), (3) that there is no gas within the tube (or at least not enough to cause retardation of the electrons due to collisions with gas molecules), (4) that equipotential electrodes are used, (5) that the electrons are emitted with zero velocity,

[1] Child, *Phys. Rev.*, 32, p. 498, 1911.

and (6) that the contact potential at the plate is negligible. Since most of these assumptions are never wholly realized in actual tubes, the exponent in Eq. (3-2) generally differs somewhat from $\frac{3}{2}$, being somewhere between 1 and $\frac{5}{2}$.

The use of filamentary-type cathodes in certain tubes is one of the reasons for variations in the exponent, especially when the plate voltage is low, due to violation of assumption (4) of the preceding paragraph. Unlike indirectly heated cathodes, the filamentary type has a difference in potential impressed across the extremities of the emitting surface equal to the filament voltage, which results in a varying potential between the anode and various parts of the cathode. This effect can best be shown by reference to Fig. 3-10, which shows filamentary-type cathode F, having 5 volts (d-c) impressed across its terminals. A 45-volt battery is used to supply the plate voltage, its negative terminal being connected to the negative terminal of the filament battery. The difference in potential between the lower end of the filament and the plate is, therefore, 45 volts, but this potential decreases along the filament until there is but 40 volts between the upper end of the filament and the plate. This varying potential between plate and cathode increases the effective exponent in Eq. (3-2) to about $\frac{5}{2}$ for plate potentials which are comparable to the filament drop. As the plate voltage is increased, the effect of the filament drop is evidently lessened, and the exponent tends to approach the theoretical $\frac{3}{2}$.

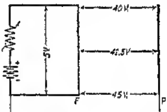


FIG. 3-10. Illustrating the variation of plate potential along a filamentary-type cathode.

Energy Loss on Plate. According to Eq. (3-2) the plate current in a tube will increase indefinitely as the plate voltage is increased, provided there is sufficient emission. Actually there is a limit to the permissible current through a tube imposed by the amount of heat that the tube will dissipate. As pointed out in Chap. 1, each electron passing through the space from cathode to anode will fall with increasing velocity toward the plate. When it strikes, it will give up most of its kinetic energy to electrons and atoms within

the plate, the velocity of which will be thereby increased; *i.e.*, the temperature of the plate will rise. The kinetic energy of each electron as it strikes the plate (neglecting velocity of emission) will be

$$\text{Kinetic energy} = E_b e \text{ joules} \tag{3-3}$$

where E_b = plate voltage

e = charge on electron, coulombs

If Eq. (3-3) is now multiplied by the number of electrons flowing per second, the total loss on the plate will be given in the familiar form of

$$P = E_b e n = E_b I_b \text{ joules/sec, or watts} \tag{3-4}$$

where n is the number of electrons flowing per second; and the resistance of the tube will be

$$\frac{E_b}{I_b} = R_b \tag{3-5}$$

The subscript b is applied to this resistance, since it is obtained under static, or d-c, conditions as distinguished from the dynamic, or a-c, resistance r_p, discussed in the next section.[1]

It should be noted that there is no actual resistance in the tube itself. In a properly evacuated tube an electron will encounter no obstruction in passing over from cathode to plate. The resistance is purely a measure of the energy required to accelerate the electrons through the space from cathode to anode, energy that reappears as heat in the plate. The rate at which this heat may be dissipated is what limits the permissible current through a tube. In very large tubes this loss becomes so great that water cooling is resorted to (see Figs. 2-11 and 2-12).

An examination of Fig. 3-3 will also show that the ratio of E_b to I_b, or R_b, is far from constant. This is characteristic of vacuum tubes where the current flowing does not vary in direct proportion to the applied potential.

Obviously the best design of a diode, other things being equal, is one that permits a given current to flow with a minimum applied plate voltage. This follows since the energy of impact of each electron as it strikes the plate, and therefore the amount of energy that the plate must dissipate in the form of heat, increases with the

[1] See Appendix A for the meaning of the various subscripts.

plate voltage. Reference to Eq. (3-1) indicates that close spacing between cathode and anode reduces the voltage required for a given current; thus tubes are designed with as close spacing as insulation requirements will permit. Since diodes are required to pass current with a positive voltage applied to the plate but must prevent the flow of current when a negative voltage is applied, they must be designed with sufficient electrode spacing to prevent breakdown under negative voltages which may be quite high.

A-C Plate Resistance r_p. Under certain conditions it may be of interest to determine the ratio of a small change in plate voltage

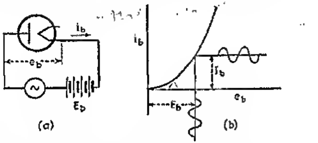


FIG. 3-11. Illustrating the significance of r_p in a diode.

to the corresponding change in plate current. Thus if a direct voltage E_b of suitable magnitude is applied to the plate of a diode, small changes in this emf will alter the current over a sensibly straight portion of the characteristic (Fig. 3-11b).[1] These changes may take the form of a small alternating potential superimposed on the direct voltage as indicated in the circuit of Fig. 3-11a. Such an alternating potential will, therefore, produce a small alternating current of nearly the same wave shape as the voltage and superimposed on the normal direct current, as shown in Fig. 3-11b.

The limit ratio of this voltage and current, as the voltage is made smaller without limit, is known as the *a-c* or *dynamic plate resistance* r_p. This ratio may be expressed mathematically as

$$r_p = \frac{de_b}{di_b} \tag{3-6}$$

and is evidently equal to the reciprocal of the slope of the characteristic curve.

[1] See Appendix A for significance of symbols used in this figure.

Actually the ratio of Eq. (3-6) varies with *any* change in applied potential no matter how small, since the slope of the characteristic curve is not constant. Thus if the magnitude of the alternating potential in Fig. 3-11 were considerably increased, the curve of alternating plate current would evidently differ in wave shape from that of the applied emf owing to curvature of the characteristic curve, and the a-c resistance would vary throughout the cycle of impressed emf. This is the obvious result of operating on a curved current-voltage characteristic.

Effect of Gas. If a tube contains an appreciable amount of gas, its behavior will be somewhat erratic. A set of characteristic curves taken on such a tube may correspond exactly to those of a normal tube up to a certain point when the plate current will often jump to much larger values than would normally be expected. The increase in current is a result of ionization of the gas, the negative ions flowing to the plate and the positive ions to the cathode, the latter neutralizing a large part of the space charge during this passage and thus permitting an increased flow of electrons through the tube.[1] It might be supposed that such an increase in current would be desirable, but unless proper precautions are taken, as in the mercury-vapor tubes, the relatively heavy positive ions will strike the cathode with such force as greatly to shorten its life. Also, a tube containing gas may sometimes ionize and pass current when *negative* voltage is applied to the plate, thereby nullifying its action as a rectifier. The presence of gas in a high-vacuum tube is therefore to be avoided.[2]

The question may well be raised as to just how high the vacuum in a tube must be before it can be said that the tube contains no gas. As a matter of fact no pump has yet been devised that will remove all the gas from a vessel. Even at a pressure of only 10^{-8} mm (a very high vacuum) there are billions of atoms of gas in every cubic centimeter of space. However, the purpose of evacuating the space through which the electron must travel is to permit the electron to make its journey from cathode to plate without

[1] A detailed explanation of this phenomenon will be found in Chap. 4, since it is characteristic of gas-filled tubes.

[2] As a matter of fact the most serious effect of gas in a *diode* is cathode destruction. Since this may be avoided by proper design, most diodes are gas-filled rather than high-vacuum. Triodes, however, offer different problems and are most commonly high-vacuum. For gas-filled diodes and their applications see Chaps. 4 and 7.

obstruction. If the tube is therefore evacuated to such a degree that it is very unlikely an electron will encounter an atom in its passage over to the plate, it may be said that there is no gas present in the tube.

To gain some conception of just how high the vacuum should be, assume a tube the anode and cathode of which are spaced 1 cm apart. Then, if an electron is to travel from cathode to anode without striking an atom of gas, it must be capable of traveling at least 1 cm without the probability of encountering a gas atom. The diameter of an electron has been determined as 4×10^{-13} cm; that of a gas atom is about 10^{-8} cm, its size depending somewhat upon the kind of gas of which the atom is a constituent part. If the vessel is now assumed to be evacuated to a pressure of 7×10^{-5} mm, leaving about 10^{12} atoms of gas per cubic centimeter, the spacing between gas atoms is about 10^{-4} cm, or nearly *a billion times the diameter of an electron*. In other words, an electron will pass between two atoms as readily as a pea would pass through a hole having a diameter equal to that of the earth! Obviously, the probability of an electron striking an atom is very remote.

This same problem can be analyzed more scientifically by a consideration of the mean free path of an electron. The *mean free path* of an electron may be defined as the probable distance that it will travel without a collision and is given by the equation

$$l = \frac{1}{\pi r^2 n}$$

where l = length of mean free path, cm

r = radius of gas atom, cm

n = atoms/cu cm

For the problem given above, this equation will give

$$l = \frac{1}{\pi \left(\frac{10^{-8}}{2}\right)^2 \times 10^{12}} = 12{,}700 \text{ cm}$$

Although this distance is only *average* and some electrons may travel a much shorter distance before striking an atom of gas, it is obvious that the number that will do so in the short distances found in the vacuum tube is negligibly small, so that a tube with the electrode spacing and evacuation just assumed would be said to be free of gas (see also page 34).

Rectifier Action. The principal use of the two-element, high-vacuum tube is that of rectification of alternating currents. It is used extensively in rectifying currents of power frequencies (generally 25 or 60 cycles) in order to secure a d-c supply and has also found application at radio frequencies.

Figure 3-12 shows a simple circuit using a diode as a rectifier. One lead from the a-c source is connected to the plate of the tube, and the d-c load (indicated by the resistance R_L) is taken off between the cathode of the tube and the other terminal of the a-c supply. During the half cycle of the a-c supply which impresses positive voltage on the plate of the diode, current will flow through

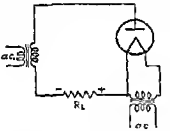


FIG 3 12 A simple diode rectifier circuit

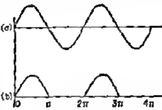


FIG 3 13 Wave shape of the current flowing in the circuit of Fig 3-12

the circuit in an amount depending upon the resistance of the load and the tube drop, but during the negative half cycle no current whatsoever will flow through the circuit. Therefore the emf appearing across the load resistance will be pulsating but unidirectional. The impressed emf is shown in Fig 3-13a; and the current flowing, in Fig 3-13b, assuming a resistive type of load.

For most purposes the pulsating current that flows through the load in the circuit of Fig. 3-12 would be entirely unsatisfactory. Actually this current may be made as constant as desired by the use of *filters* or polyphase a-c circuits or both. These are discussed in Chap. 7 for low-frequency applications and in Chap. 14 for r-f applications.[1]

2. TRIODES

In 1907, Lee DeForest inserted a third element, or grid, between the plate and filament of the Fleming valve and laid the founda-

[1] In r-f service the tube is commonly known as a *detector* or *demodulator*.

tion for some of the most notable achievements of the age. Radio broadcast and reception, the long-distance telephone, facsimile-picture transmission, public-address systems, automatic and remote control of all sorts of power machinery, television, radar, d-c power transmission, and a nearly unending list of other achievements have been, or are being, made possible by the invention and further development of the three-element tube or triode, together with the diode and the more recent multielement tubes.

The primary purpose of this third element in a vacuum tube is to control the flow of plate current by the insertion of an electric charge between anode and cathode which will either increase or decrease the effect of space charge. In the ideal case this charge should be inserted without (1) in any way retarding the current flow mechanically or (2) absorbing any current. The first condition is approached by constructing the grid of a mesh of wires in order to leave as much space as possible through which the electrons may pass on their way to the plate.[1] The second condition may be approached by maintaining the grid at a potential more negative than the cathode, which will tend to prevent any electrons from being attracted to the grid.

Static Characteristic Curves. There are three independent variables in a high-vacuum triode: cathode temperature, plate potential, and grid potential; and two dependent variables: plate current and grid current. Curves may be obtained by permitting any one of the first three to vary while the other two are held constant. Such curves are termed the *static characteristic curves* of the tube. For most purposes, however, cathode temperature is not considered a variable, since, in a large majority of applications, it is necessary only that the total emission be at least several times the normal plate current. The cathode heater or filament voltage specified by the manufacturer is generally used in obtaining the characteristic curves

Figure 3-14 shows a circuit for obtaining these curves for a triode. The grid battery (or other d-c source) is divided into two parts; thus the grid may be made either positive or negative with respect to the cathode. A zero-center voltmeter is desirable for measuring e_c. Note that all potentials are measured from the cathode; as all potentials are relative, it is necessary that some point be used as a reference from which to measure. For filamen-

[1] Details of tube construction were given in Chap. 2.

tary-type cathodes the reference point is the negative filament terminal when direct current is used to heat the cathode or the mid-tap of the transformer when alternating current is used (as in Fig. 7-6). Henceforth when the potential of the grid or plate

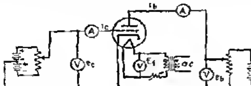


FIG. 3-14 Circuit for determining triode tube characteristics.

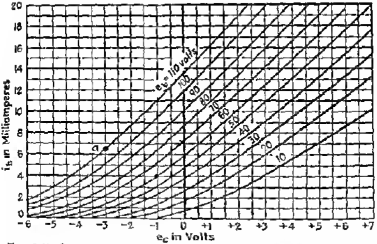


FIG 3-15 Static characteristic curves of a triode. Plate current vs grid voltage

is stated, it should be understood that the potential is measured between the grid or plate and the appropriate reference point at the cathode.

Figure 3-15 shows a set of characteristic curves of a triode. Note that each curve is similar in shape to curve 1, Fig. 3-3, where the plate current was limited by space charge. In Fig. 3-16, curves of i_b vs. e_b are plotted by cross-reading the curves of Fig

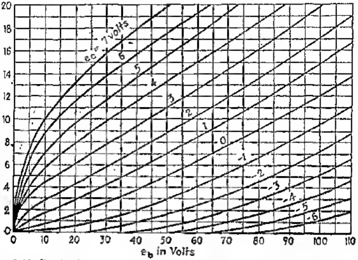


FIG. 3-16. Static characteristic curves of a triode. Plate current vs. plate voltage.

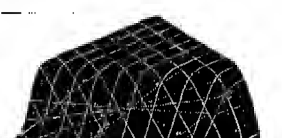


FIG. 3-17. Characteristic surface of a triode vacuum tube.

3-15. By combining these two sets of curves, a characteristic surface[1] is obtained such as the one pictured in Fig. 3-17 where the

[1] J. R. Tolmie, The Characteristic Surfaces of the Triode, *Proc. IRE*, 12, p. 177, April, 1924; Characteristic Surfaces of Thermionic Valves, *Univ. Wash. Eng. Exp. Sta. Bull.* 24, 1924.

plate and grid voltages are measured along the two horizontal axes and the plate current on the vertical axis. The lower horizontal plane represents zero plate current; and the higher plane, saturation. The normal operating range lies on the steep slope between these two planes. The current in Figs. 3-15 and 3-16 was not carried to sufficiently high values to produce saturation since we are not normally interested in currents of that magnitude.

At this point it is well to note again that the static characteristic curves of a vacuum tube accurately portray the *instantaneous* action of the tube under *any* conditions whatsoever, except when the transit time of the electron becomes important, as at frequencies of the order of a few hundred million and more. Within these limits, no matter what the external circuit may be, if the *instantaneous* values of any two of the three variables i_b, e_b, and e_c are known, the third may *always* be found for *that instant* by reading from the static characteristic curves. The same thing is true of the three variables i_c, e_b, and e_c. The student should bear this carefully in mind as he studies the following pages.

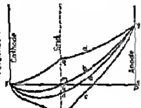


FIG. 3-18. Potential distribution curves for a triode. Curve *a* is for a "free" grid.

Potential Distribution. As in the case of the diode, the characteristic curves of the triode may best be explained by a study of the potential distribution within the tube. Let curve *a*, Fig. 3-18, represent the distribution of potential inside the tube with the grid removed. It is assumed that the cathode temperature is normal and is therefore sufficiently high to provide an excess of electrons under any applied plate potential within the normal range of the tube, the shape of this curve must therefore be similar to curve *d*, Fig. 3-5, and not to curve *b*, where saturation current is flowing.

Let the grid now be inserted at point G and let it remain entirely disconnected electrically from any portion of the circuit and from ground. The potential existing at the point *G*, before insertion of the grid, was negative and equal to *mn*, the grid will therefore attract electrons until it assumes the potential *mn*, when equilibrium will exist. Since the grid is entirely disconnected from any external circuit, or free, this potential is known as the *free-grid potential*.

toward plate, whereas the positive direction of flow of i_p is considered as being from plate toward load. The logic of these assumptions will be discussed shortly.

Equation (3-18) contains both alternating and direct components. The alternating components are $-i_p(r_p + R_L)$ and μe_g; the direct components are $I_b(r_p + R_L)$ and $\mu E_c + E_{bb} + r_p c$ (μ and r_p being assumed constant). If the quantities on either side of the equality sign in Eq. (3-18) are to be identical, as the equality sign indicates, the alternating and direct components of each side of the equation must be, respectively, identical. We may then equate the alternating components on each side of the equation, giving

$$-\mu e_g = i_p(r_p + R_L) \tag{3-19}*$$

or, in rms values

$$-\mu E_g = I_p(r_p + R_L) \tag{3-20}*$$

This equation is the basis for all design of class A amplifiers and is used quite widely in the solution of other types of circuits. *It is probably the most important equation in vacuum-tube engineering.* Nevertheless it is well to point out again that it was obtained by assuming r_p and μ to be constant[1] and *may be applied only when these tube coefficients are constant* or practically so and, therefore, only when the grid and plate potentials are such as to confine operation to the reasonably straight portions of the static characteristic curves.[2] Fortunately such is the case in a very large number of vacuum-tube applications.

The logic of assuming $i_b = I_b - i_p$ rather than $i_b = I_b + i_p$ may now be seen. It is evident from Eq. (3-19) that in so far as the alternating components are concerned, the tube acts like an alternator with an instantaneous induced emf of $-\mu e_g$ and an internal impedance r_p, Fig. 3-24. The *apparent* source of energy from which i_p flows is therefore within the tube, and the positive direction of flow of i_p is assumed to be *out of the plate*. On the other hand, the source of energy from which I_b comes is the plate

[1] This assumption was, of course, implied by the partial differentiations in Eqs. (3-10), (3-11), and (3-12).

[2] It may sometimes be used where these stipulations do not apply, to give equivalent sine-wave conditions of sufficient accuracy for the purpose at hand.

battery E_{bb}, therefore, the positive direction of this current is assumed to be *from* the positive terminal of the battery *to* the plate. Hence the minus sign between I_b and i_p when they are combined to give i_b.[1]

The terminal voltage of the equivalent alternator of Fig. 3-24 appears across the load resistance R_L and is, of course, I_pR_L, or E_p. Equation (3-20) may therefore be written

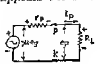


FIG. 3-24 Equivalent circuit of a triode amplifier

$$-\mu E_g = I_pr_p + E_p \tag{3-21}$$

or

$$\mu E_g + E_p = -I_pr_p \tag{3-22}$$

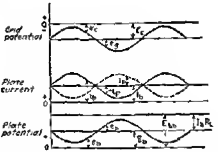


FIG. 3-25 Instantaneous plate current, and plate and grid voltages, in a triode amplifier with a sine-wave impressed grid voltage. Resistive load

Phase Relations in an Amplifier. A graphical representation of the voltages and currents in a class A triode amplifier with resistance load is shown in Fig. 3-25. The grid voltage is assumed to be a sine wave of emf superimposed on the grid bias; the plate current is found by means of the method of Fig. 3-21 or with the aid of Eq. (3-19); and the plate voltage is found from Fig. 3-21 or from Eq. (3-15). The total instantaneous plate voltage is always less than the plate-battery potential by the instantaneous drop through the load circuit, so that, as the total current increases, the potential must necessarily decrease. This is clearly indicated by the curves of i_b and e_b, Fig. 3-25. Since $i_b = I_b - i_p$, the difference between the i_b curve and the direct current I_b is indicated as $-i_p$; to show the actual phase relations between the al-

[1] See Appendix E for further treatment of this problem.

ternating components, the positive alternating plate current is also shown by a dotted curve. This is seen to be in phase with the alternating component of plate voltage as it should be for a resistance load.

A vector diagram of the voltages and currents involved is shown in Fig. 3-26.[1] I_p is the reference vector; $I_p r_p$ and E_p are drawn in phase with this current. $-\mu E_g$ is equal to $(I_p r_p + E_p)$ and is therefore also drawn in phase with I_p. This diagram, except for the E_g vector, is exactly similar to standard alternator vector diagrams found in texts on alternating currents, except that the internal impedance is in this case resistive only.

FIG. 3-26. Vector diagram of a triode amplifier with resistive load.

The load circuit of a vacuum tube may not consist of pure resistance; it may contain either inductive reactance or capacitive reactance (Fig. 3-27). It is obvious that a condenser cannot be used directly in series with the plate circuit, since it would prevent

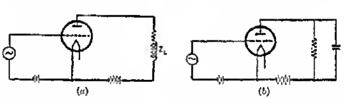


FIG. 3-27. Simple circuit of a triode amplifier, (a) with inductive load and (b) with capacitive load.

the flow of direct current, making the tube inoperative. However, it may frequently be found in parallel with an inductance or, as in Fig. 3-27b, with a resistance. When inductance is used, the load circuit may present condensive reactance, inductive reactance, or pure resistance, depending upon whether the frequency of the input signal is, respectively, greater than, less than, or equal to the resonant frequency of the load. This is a condition frequently encountered when the tube is used at radio frequencies.

Equation (3-20) may be altered to fit the more general case of inductive load by replacing R_L with Z_L

[1] See Appendix E for the methods of indicating the relative directions of alternating currents and voltages.

$$-\mu E_g = I_p(r_p + Z_L) \qquad (3\text{-}23)$$

where $(r_p + Z_L)$ must be treated as a complex quantity.

Figure 3-28 is a vector diagram for an inductive load. I_p is the reference vector, I_pr_p is in phase with I_p; and E_p is leading by an angle θ, where θ is the phase angle of the load circuit. $-\mu E_g$ is, of course, the vector sum of E_p and I_pr_p. It will be seen that none of the three variables I_p, E_p, and E_g is in phase, I_p lagging $-\mu E_g$ by the angle ϕ and E_p by the angle θ where

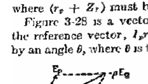


FIG. 3-28. Vector diagram of a triode amplifier with inductive load

$$\phi = \tan^{-1}\frac{\omega L}{r_p + R_L} \qquad \theta = \tan^{-1}\frac{\omega L}{R_L}$$

where L = inductance of load circuit

R_L = resistance of load circuit

If ωL is made increasingly large, as compared with R_L and r_p, the two voltage vectors will approach a 90-deg phase position from the current.

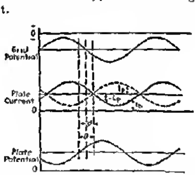


FIG. 3-29. Instantaneous plate current, and plate and grid voltages, in a triode amplifier with a sine-wave impressed grid voltage. Inductive load

The relative phases are also clearly shown in Fig. 3-29.

Figure 3-30 is a vector diagram for capacitive load, and Fig. 3-31 shows the instantaneous current and voltages for this type of load.

Vacuum-tube Coefficients. The amplification factor, plate resistance, and mutual conductance (grid-plate transconductance)

are often referred to as the coefficients of a tube. As already indicated they are, to a considerable extent, functions of the instantaneous voltages applied to the electrodes. The mathematical expressions for these coefficients have been given previously but are repeated below.[1]

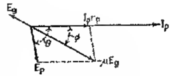


FIG. 3-30. Vector diagram of a triode amplifier with capacitive load.

$$\mu = -\frac{\partial e_b}{\partial e_c} = -\frac{\partial e_p}{\partial e_g} \qquad (3\text{-}24a)^*$$

$$r_p = \frac{\partial e_b}{\partial i_b} = -\frac{\partial e_p}{\partial i_p} \qquad (3\text{-}24b)^*$$

$$g_m = \frac{\partial i_b}{\partial e_c} = -\frac{\partial i_p}{\partial e_g} \qquad (3\text{-}24c)^*$$

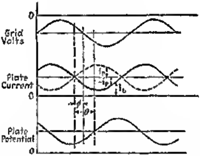


FIG. 3-31. Instantaneous plate current, and plate and grid voltages, in a triode amplifier with a sine-wave impressed grid voltage. Capacitive load.

The reader should again note that r_p is the *a-c plate resistance*, as distinguished from the d-c plate resistance R_b, which is merely E_b/I_b (see also page 46). The product $I_p r_p$ of Eqs. (3-21) and (3-22) gives the internal tube drop due to the alternating component alone.

[1] The second expression for each term is derived from the first by substituting the relations $e_c = E_c + e_g$, $e_b = E_b + e_p$, and $i_b = I_b - i_p$ where E_c, E_b, and I_b are, of course, constant.

Figure 3-32 shows curves of μ, r_p, and g_m of a triode. The independent variable is i_b rather than e_b or e_c, since μ, r_p, and g_m are relatively independent of changes in the plate and grid potentials, provided both are varied in such a manner as to keep the plate current constant. This statement is borne out by the curves of Figs. 3-15 and 3-16 where it may be seen that the slope of all curves is nearly the same for a given value of i_b.

Grid Current. Little has been said so far about the possibility of current flow in the grid circuit. The fundamental purpose of inserting a grid in the tube is merely to alter the effect of the space charge and so control the flow of plate current, yet electrons will

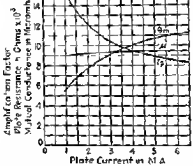


FIG. 3-32. Curves of μ, r_p, and g_m of a triode.

flow to the grid just as to the plate, if it is maintained at a potential more positive than the free-grid potential. In Fig. 3-33 are shown curves of grid current in a triode determined by means of the circuit of Fig. 3-14. The grid current is seen to vary exponentially with grid voltage just as did the plate current. All curves necessarily come to zero at the free-grid potential (practically at $e_c = 0$ for this tube), defined as that potential at which electrons just cease to flow to the grid.[1] This point varies slightly for each

[1] Actually this definition is not quite rigid, since some positive ions are always present. The free-grid potential is the potential at which equilibrium exists between the flow of electrons and of ions to the grid, the net grid current being zero. In a highly evacuated tube the number of ions is, of course, small.

curve, although the change is too small to be noticeable in the figure.

As the plate voltage is increased, the grid current is seen to *decrease*, an effect shown more clearly in Fig. 3-34. That this should have been expected may be seen from the potential distribution curves (Fig. 3-35). These curves all represent the distribution with free grid, the free-grid potentials for the plate

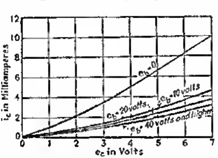


FIG. 3-33. Static characteristic curves of a triode. Grid current vs. grid voltage.

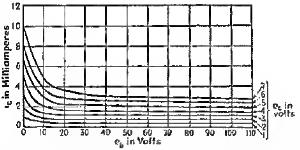


FIG. 3-34. Static characteristic curves of a triode. Grid current vs. plate voltage.

voltages a, b, and c being p_1, p_2, and p_3, respectively. If the grid potential is now set at p, current will flow to the grid under the potential pp_1 with plate voltage a. Similarly for the lower plate voltage b, the grid current is flowing under a potential pp_2, greater than for a. For c, the effective grid potential pp_3 is still greater.

This may also be stated by saying that an electron traveling toward the interstices of the grid, as x in Fig. 3-36, may be drawn to one side and so into a grid wire if the plate voltage is low, but,

if the plate voltage is increased, the effect of the grid on this electron may be overpowered, and the electron go to the plate instead. Electrons traveling directly toward the grid wires, as y, will, of course, go into the grid, regardless of the magnitude of the plate voltage, as long as the grid remains positive. Therefore the grid current will be larger at the lower plate potentials.

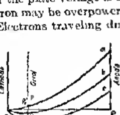


FIG. 3-35. Potential distribution curves illustrating the effect of the plate potential on the grid current.

A flow of grid current is usually undesirable. In the circuit of Fig. 3-22 it will have the effect of requiring that additional power be delivered by the a-c source. Also, if the input alternator has internal impedance, the flow of grid current will produce a voltage drop which will be greater on the positive half cycle than on the negative half cycle so that a distorted wave will be applied to the grid. This phase of the problem will be discussed more fully later (page 351).

3. MULTIGRID TUBES

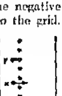


FIG. 3-36. Electron x will go to the plate at high plate potentials but may go to the grid when the plate potential is low. Electron y will go to the grid regardless of the plate potential.

In a vacuum-tube amplifier the current in one circuit (plate) is controlled by the potential applied to another (grid). This process should preferably be unilateral, *i.e.*, the current and potential set up in the plate circuit should have no effect on the grid potential. Since electrons flow only from the cathode through the grid structure to the plate (and not in the reverse direction), the operation of the tube would be entirely unilateral if it were not for the electrostatic capacitance between plate and grid. As is more fully demonstrated on page 424, this capacitance may have very undesirable effects, especially at radio frequencies.

Tetrodes. Screen-grid Tubes. The screen-grid tube was developed primarily to eliminate the grid-to-plate capacitance, although the applications of this tube have extended considerably beyond this original purpose. An additional grid, or screen, is inserted between the regular control

grid and the anode (Fig. 3-37).[1] Shielding is also supplied on the outside of the plate. In glass-envelope tubes (Fig. 3-37a) this is done by slipping a metal shield can over the tube after it has been inserted in its socket, and even better results are obtained with metal-envelope tubes (Fig. 3-37b) by grounding the envelope. An additional shield at the top of the tube surrounds the grid lead to eliminate capacitance between this lead and the plate. In glass-envelope tubes the shield can should have a metal ring that fits closely around the outside of the glass opposite this internal shield as shown.

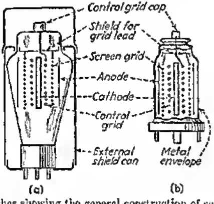


FIG. 3-37. Sketches showing the general construction of screen-grid tubes, especially the methods of providing adequate shielding. Present-day practice is to use pentodes which include a third or suppressor grid (see Fig. 3-51).

When this tube is used as an amplifier, the screen grid is tied to the cathode *dynamically* so that no change in potential can occur between these two elements, thus providing an effective electrostatic shield between plate and grid. *Statically* the screen grid is maintained at a positive potential relative to the cathode to permit a normal flow of electrons through the tube. These two conditions are secured in practice by returning the screen-grid lead to a point of positive potential and connecting a low-impedance condenser between screen grid and cathode.

[1] Tubes of this type are now built with *three* grids, not two, as described on p. 77, but the general construction is as shown, especially with regard to the shielding.

The mechanism of the shielding action may be seen from Fig. 3-38. A triode is represented in (*a*) of Fig. 3-38, and a screen-grid tube in (*b*), with only the alternating voltages shown in the external circuits. It is evident that in the case of the triode the

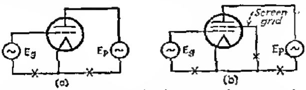


FIG. 3-38 Illustrating the shielding action of a screen grid.

electrostatic field set up by the plate voltage must extend to the cathode, thereby inducing a voltage on the grid. In the screen-grid tube, however, the alternating plate voltage is set up between anode and screen, and the region between screen and cathode, in which the regular (or control) grid is located, will have zero field in so far as the plate voltage is concerned. The fact that direct voltages must also be included in a practical circuit (at the points indicated by crosses) in no way affects the a-c fields. The d-c fields are of course of no interest, since they are steady and will cause no alternating current to flow in the grid circuit.

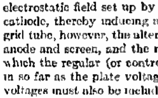


FIG. 3-39 Curves of total space current in a screen grid tetrode for various screen grid voltages, plotted against control-grid voltage.

Tetrode Static Characteristic Curves. If curves of total space current $i_b + i_{c2}$ (where i_{c2} is the instantaneous current flow to the screen grid) are plotted against control-grid voltage e_{c1} for various *screen-grid* voltages, the resulting curves (Fig. 3-39) will be very similar to the plate-current–grid-voltage curves of the triode (Fig. 3-15, page 52), where the *plate* voltage was changed in steps. In other words, in so far as the space current flowing away from the cathode is concerned, the screen grid functions as an anode so that cathode, control grid, and screen grid constitute a more or less conventional triode in themselves. It is interesting to note that this space current is, however, virtually independent of the *plate*

voltage, since the screen grid serves as an electrostatic shield to prevent the plate from affecting the field around the cathode. If this shielding action is perfect, the plate can have no effect whatsoever on this field (space charge) and can, therefore, cause no change in the number of electrons leaving the vicinity of the cathode. In practice the shielding, although not perfect, can be made sufficiently effective to prevent the plate from causing more than a few per cent change in space current when its potential is varied from zero to its maximum safe value.

In Fig. 3-40 curves of i_b, i_{c2}, and of their total are plotted against plate voltage for three different control-grid voltages. The curves

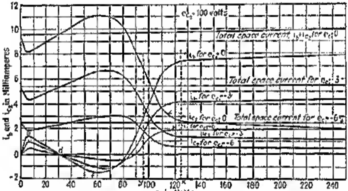


FIG. 3-40. Curves of plate current, screen-grid current, and total space current in a tetrode, plotted against plate voltage.

of total space current are virtually flat, bearing out the statement already made that this current is nearly independent of the plate voltage. On the other hand, the plate current is far from independent, unless the plate potential is somewhat greater than that of the screen grid. The space current divides between the screen grid and plate in a manner determined in part by the potential applied to each, their relative areas, and the distance of each from the cathode, although, if these were the only factors involved, the plate current would rise rather abruptly to its final value and would not have the negative dip shown. Unfortunately, as each electron from the cathode strikes the screen grid or plate, it tends to knock out one or more additional electrons, a phenomenon described in Chap. 1 as secondary emission. These secondary electrons tend to

travel to the electrode of highest potential and so constitute a flow of current from the one electrode to the other.

In Fig. 3-40, as the plate potential is increased from zero, the plate first attracts an increasingly larger percentage of the electrons being drawn away from the cathode by the screen grid, so that the plate current rises, while the screen-grid current decreases by a nearly equal amount. As the plate voltage is still further increased, the velocity of impact of the electrons becomes sufficient to produce a copious secondary emission; and since the screen-grid potential is still higher than that of the plate, these electrons are drawn over to the screen, thereby increasing the screen-grid current and decreasing the plate current. At one point in the curves of Fig. 3-40 the emission of secondary electrons becomes so profuse that they exceed in number those flowing *to* the plate, and the net plate current is negative.[1] Beyond this point, although secondary emission continues to increase, electrons so emitted tend to fall back into the plate, its potential being virtually as high as that of the screen grid. Beyond this point, too, the plate soon begins to attract secondary-emission electrons from the screen grid, although the number of such electrons is small—*never sufficient* to produce a negative screen-grid current in commercial tubes.

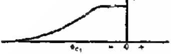


FIG. 3-41. Dynamic i_b-e_{c1} curve for a tetrode in which the minimum instantaneous plate voltage is less than the screen-grid voltage.

The secondary-emission effect just described limits the range throughout which the tube may be operated. If the plate voltage is allowed to swing below the potential x, the dynamic i_b-e_{c1} curve will have a serious bend at the upper end, such as in Fig. 3-41, where the minimum plate voltage is about at y (Fig. 3-40). For example, the normal plate voltage on a certain screen-grid tube is 180 volts, and the plate current is virtually independent of the plate voltage down to about 100 volts. The maximum alternating voltage permissible in the plate circuit is then

$$(180 - 100)0.707 = 56.5 \text{ volts rms}$$

whereas, if the plate current were independent of the plate voltage

[1] These curves were taken on an early type of tube, chosen to emphasize the effect of secondary emission.

to a value as low as, say, 20 volts, the alternating voltage could be double, or 113 volts.

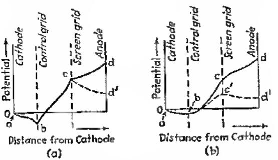


FIG. 3-42. Potential distribution curves of a screen-grid tetrode.

Potential Distribution. Curves of potential distribution within the screen-grid tube may be drawn similar to those of Fig. 3-18 for the triode. Four such curves are shown in Fig. 3-42 for different anode voltages, other electrode potentials being held constant. The curves shown in (a) are taken along a line intersecting wires of the screen and control grids, as AA' in Fig. 3-43. It will be seen that the portion of the curve *abc* is quite similar to curve *c* of Fig. 3-18 for the triode and is entirely independent of the potential distribution from screen to anode as indicated by the two curves *abcd* and *abcd'*, curves *cd* and *cd'* representing two different anode voltages, other electrode potentials being constant. Evidently the number of electrons traveling through the region between cathode and screen will be the same regardless of the anode voltage; *i.e.*, the space charge is unaffected by the anode.[1]

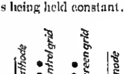


FIG. 3-43. Sketch showing two possible paths in a screen-grid tube along which the potential distribution may be considered.

On the other hand if a path is considered that does not intersect a wire of the screen, as BB', Fig. 3-43 (where the path shown also

[1] Although electrons cannot actually travel along a path intersecting a grid wire, these potential distribution curves apply with reasonable accuracy to a path passing close to a grid wire.

passes between wires of the control grid although this is not a necessary condition), the potential curves will be as in (*b*) of Fig. 3-42, where the curves *abcd* and *abc'd'* represent two different plate potentials, all other electrode potentials remaining unchanged. In this case the potential at the plane of the screen grid will obviously be affected by the anode, the relative effect of anode and screen grid being a function of their respective potentials and of the proximity and area of their nearest surfaces. With such a potential distribution it is obvious that the space current between

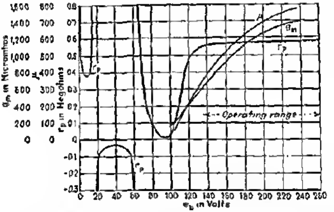


FIG. 3-44 Curves of μ, r_p, and g_m for a typical screen-grid tetrode

cathode and screen will be somewhat affected by the potential on the anode. Thus perfect screening requires the insertion of a continuous screening surface across the entire electron path, which would, unfortunately, effectively prevent the passage of any electrons through to the anode. Perfect shielding in this type of tube is therefore impossible.

Variations in μ, r_p, and g_m with e_b. Figure 3-44 shows how μ, r_p, and g_m vary with the plate voltage of a typical screen-grid tube. r_p is negative for quite a range of plate voltage, because of the negative slope of the plate-current-plate-voltage curve, a characteristic that makes many screen-grid tubes suitable as dynatron oscillators (see page 477). Throughout the operating range of the tube the plate resistance is extremely high,[1] its magnitude

[1] The operating range for this tube lies above about 120 volts.

being an excellent indication of the screening effect of the grid. This follows since a high resistance means that the slope of the plate-current–plate-voltage curve must be nearly horizontal or that the plate current is virtually independent of the plate voltage, the natural result of good screening.

The amplification factor μ is also quite variable with plate voltage, attaining high values at high plate voltages. A high μ is also an indication of good screening, since μ is a measure of the relative effectiveness of the plate and the grid in controlling the flow of current through the tube. If the current in the plate were entirely independent of the plate voltage, μ would be infinite.

Applications. Tetrodes are seldom used now, having been largely superseded by pentodes and beam tubes. Nevertheless an understanding of the performance and limitations of tetrodes, as given in the preceding paragraphs, is desirable before studying the more recent types of tubes now in use.

FIG. 3-45. The heavy curve is the normal characteristic of a screen-grid tetrode. The dotted curve shows the result of eliminating the flow of secondary emission current from the plate.

Pentodes. The plate-current–plate-voltage characteristic of a screen-grid tube is reproduced in the solid curve of Fig. 3-45, showing the characteristic dip caused by secondary-emission electrons traveling from the plate to the screen grid. If a means is provided to prevent the flow of these electrons, the curve will have the general shape shown by the dotted line, rising almost immediately to a value only slightly less than the total space current drawn over from the cathode by the screen grid. It is the function of the third grid in the pentode tube to produce this type of characteristic by preventing the flow of secondary electrons.

The first or inner grid of a pentode tube (the grid nearest the cathode) is the *control* grid and serves the same function as in a triode. The next one is generally known as the *screen* grid. The outer grid is the *cathode*, or *suppressor*, grid and is ordinarily tied directly to the cathode, either inside the tube or by an external connection. Figures 2-1, 2-3, and 2-4 are views of low-power, r-f pentodes.[1] Low-frequency pentodes are similar but have all leads

[1] The abbreviations r-f and a-f are commonly used for radio-frequency and audio-frequency. Audio frequencies are those which are audible to the human ear and are sometimes referred to in this book as low frequencies.

brought out at the base (as do many designs of r-f pentodes). Figure 3-46 is a sectional view of a type of r-f, high-power pentode in which the three grids are clearly distinguishable, and Fig. 3-47 shows, by way of contrast, a miniature r-f pentode for use up to 400 Mc.

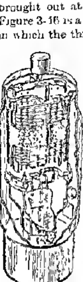


FIG. 3-46. A cut-away view of a transmitting type pentode. The plate lead is brought out to the cap, all others to pins in the base. Note the small coil springs to keep the filament from sagging, and the rigid mounting of all electrodes.

Electrons produced at the plate of a pentode by secondary emission are prevented from going over to the screen grid by the negative gradient set up by the zero-potential suppressor grid, being emitted with insufficient energy to permit them to cross this barrier. It should be noted, however, that the presence of the suppressor grid does not prevent secondary emission but merely compels all electrons emitted in this manner to return to the plate. Thus the action is similar to the triode, where secondary emission also occurs but has no effect upon the operation of the tube, since there is no positive potential point to which the electrons can be attracted other than back to the plate.

Naming the second grid the *screen grid* is somewhat misleading but seems to be standard practice. Actually the suppressor grid, in addition to its duties of suppressing the secondary emission current, is as much of an electrostatic shield as is the screen. The function of the screen grid differs from that of the suppressor in that it is raised to a positive potential to attract electrons from the region of the cathode and thus produce current flow through the tube. Without such a positive potential between the suppressor and cathode the space charge would prevent all but a very few electrons from traveling far from the cathode.

Pentode Characteristic Curves. The characteristic curves of a typical pentode are shown in Fig. 3-48. Comparison with Fig. 3-40 shows the complete elimination of the negative-dip characteristic of tetrodes. The curves are seen to be nearly horizontal over a wide range of plate potentials (the normal direct plate voltage of this tube is about 50 per cent higher than that of the tube of Fig. 3-40); and since the plate resistance of a tube is equal

to the reciprocal of the slope of the i_b-e_b curve [Eq. (3-24*b*)], it is evident that r_p is very high and much more constant than in a tetrode. The plate resistance of pentodes intended for low-power, r-f amplification is of the order of 1,500,000 ohms or higher; that of a-f, power-output pentodes is of the order of 100,000 ohms (as in Fig. 3-48). The amplification factor of the r-f type is of the order of 1500 or more; and that of the a-f type is 100 to 250. μ is more constant than in the corresponding tetrodes, but the transconductance varies in about the same manner.

FIG. 3-47. Miniature r-f pentode useful as an amplifier up to 400 Mc. (*RCA*.)

Potential Distribution. The potential distribution within a pentode tube is illustrated in Fig. 3-49. Curve *abcde* represents the potential along a line intersecting a wire on each of the three grids, as AA' of Fig. 3-50; curve *ab'c'd'e* represents the distribution along any line passing between wires on each of the three grids, as BB' of Fig. 3-50. The suppressor is assumed to be tied to the cathode in both cases. It should be obvious from inspection of the figure that, barring the effect of any collisions encountered along the way,[1] all electrons that pass the control grid will have sufficient energy to pass the suppressor grid. Some will, of course, strike the wires of the screen or suppressor grids and so return to the cathode without reaching the anode, but all others will reach the anode.[2]

A clearer picture may be drawn by considering the velocity of the electrons at various points in the tube. To illustrate, consider an electron that is emitted from the cathode with just sufficient velocity to carry it past the control grid. Its velocity at any point beyond will be a function of the difference in potential between that point and the potential b' at which the electron passed through the control grid with virtually zero velocity. This may be

[1] Collisions in a high-vacuum tube are, of course, highly unlikely.

[2] See footnote on p. 75.

seen by applying the principle of conservation of energy. As an electron passes, or "falls," from one region to another through

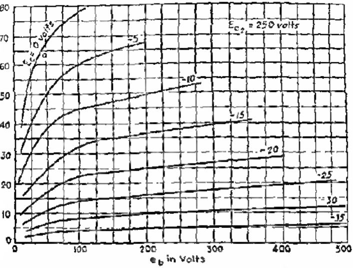


FIG 3-48. $i_b - e_b$ static characteristic curves of an a-f power-output pentode tube

a difference of potential, it loses potential energy. This loss must exactly equal the kinetic energy of motion that it gains, or

$$Ee = \frac{mv^2}{2}$$

where E = potential difference through which electron falls, volts
e = charge on electron, coulombs
m = mass of electron, kg
v = velocity of electron, m/sec

Since the only variable quantities are v and E, the premise is proved.

It is now possible to determine whether or not electrons will still have sufficient velocity when they reach the plane of the suppressor to pass on through to the anode. Inspection of Fig. 3-49 shows that the potential at all points in the plane of the suppressor grid but between its wires is slightly higher than that of the grid

structure itself, as point d'. Therefore, all points in the plane of the suppressor are at a more positive potential than points in the plane of the control grid such as b'. Thus all electrons passing the control grid and not striking the screen will reach the plane of the suppressor grid with some remaining velocity.[1] A few will, of course, strike the grid wires of the suppressor and so return to the cathode, but the majority will pass on over to the anode.

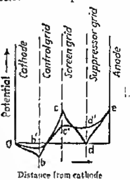


FIG. 3-49. Potential distribution curves of a pentode tube.

If the suppressor is to prevent secondary-emission electrons from passing from anode to screen grid, it is obvious that the potential d' must be sufficiently negative with respect to the anode *at the time when the anode is least positive* to stop the highest velocity secondary-emission electrons that are being emitted. Since the pentode was designed to permit large variations in plate potential, resulting in very low anode voltages at certain points in the a-c cycle, it is essential that the maximum potential in the interstices of the suppressor grid be kept as low as reasonably possible. This can be done in part by using a fine-mesh grid for the suppressor, but it must be remembered that any step in that direction will necessarily reduce the area through which electrons may flow to the anode. The effectiveness of the suppressor may also be increased by making it somewhat negative with respect to the cathode, provided the average potential in the plane of the suppressor is far enough above that at the control grid to allow electrons to reach the anode in sufficient numbers. A better solution to the suppressor problem is the beam tube described in a later section of this chapter.

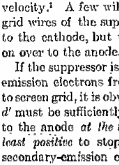


FIG. 3-50. Sketch showing two possible paths in a pentode tube along which the potential distribution may be considered.

[1] It is, of course, possible for some points in the plane of the control grid to be at a slightly positive potential if the grid wires are far apart. Even so, since all electrons must start at the cathode, they will have sufficient energy to pass the suppressor unless it is made *negative* with respect to the cathode.

Remote Cutoff Tubes. Remote cutoff tubes (also known as *variable-μ* tubes) are pentodes so designed that their coefficients (μ, g_m, and r_p) vary only gradually with changes in grid bias.[1] Such a design provides a tube for use in an amplifier the gain of which may be varied by adjustment of the direct grid voltage or bias, a method of control that is essential where automatic volume control is desired, as, for example, in radio receivers (see page 587)

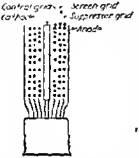


FIG. 3-51. Sketch showing the construction of one type of remote cutoff pentode

One method of designing a pentode tube of this type is shown in Fig. 3-51, which is essentially a cross-section view of Fig. 2-2, *H* (except for the shield at the top of the tube). The mesh of the control grid is varied throughout its length, being close together at the top and bottom and far apart in the center. Such a construction enables the grid to stop the flow of current entirely through the upper and lower portions of its structure at a given negative potential while still permitting current to flow through the center portion. The characteristic curve of such a tube is shown in Fig. 3-52, along with the curve of the conventional sharp cutoff type of tube. It will be seen that the curvature with the remote cutoff tube is much less than with the other, therefore, the change in tube constants will be much more gradual as the bias is varied. Unless such change is gradual, the performance of the tube as an amplifier may be seriously impaired.[1] On the other hand, a greater change in voltage is required to control the volume, but the increase is not so great as to be prohibitive.

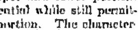


FIG. 3-52. Sketch showing the i_b-e_{c_1} curve of a conventional sharp cutoff pentode together with one of a remote cutoff pentode

Beam Power Tube. The principal limitation to the power output of the pentode tube lies in the magnitude of the third-harmonic distortion component generated. This distortion is

[1] Amplitude distortion is one undesirable result (see p. 258); another is cross modulation (p. 531).

caused primarily by the curvature of the i_b-e_b characteristic at low plate voltages, as in the vicinity of a, Fig. 3-48. An investigation led to the conclusion that this is largely due to the use of a grid type of construction for the suppressor, resulting in nonuniform suppressor action across the electron path.

The ideal suppressor is evidently one having a uniform effect over the entire area of plate-current flow. Such an ideal cannot possibly be achieved by a mechanical grid structure (see discussion of Fig. 3-49) but is possible if a zero-gradient space charge can be set up between plate and screen grid. Such a charge may exist to some extent between any two electrodes of a tube when current is flowing, especially when electrons flow from the higher to the lower potential, as in a screen-grid tube with low plate voltage. Under such circumstances the electrons slow down and form a low-gradient charge near the electrode of lower potential (plate). If the electron density is sufficiently great and is uniform throughout the path of electronic flow, a virtual cathode will be formed. (A virtual cathode is a region at approximately zero potential and having low electron velocities and high electron densities.)

The design of the beam tube is such as to make use of the virtual cathode method of suppressing the flow of secondary-emission electrons from the plate.[1] The very high density of electrons required is obtained through confinement of the electrons to beams by means of beam-forming plates located on either side of the tube and by using the same spacing between wires on the control grid and on the screen grid, with the screen grid being so mounted as to lie in the electronic "shadow" of the control grid. Thus an electron that has passed through the control grid may continue on to the plate without striking the screen or being appreciably deflected from its path. This type of grid construction also decreases the screen-grid current, since no electron can strike the screen except by following a curved path around a wire of the control grid.

A drawing of this tube is shown in Fig. 3-53b, clearly indicating the electron beams formed by the grid structure and the two beam-forming plates. The latter are connected electrically to the cathode and, in addition to confining the electrons to beams,

[1] R. S. Burnap, New Developments in Audio Power Tubes, *RCA Rev.*, 1, p. 101, July, 1936; J. F. Dreyer, The Beam Power-output Tube, *Electronics*, 9, p. 18, April, 1936; O. H. Schade, Beam Power Tubes, *Proc. IRE*, 26, p. 137, February, 1938.

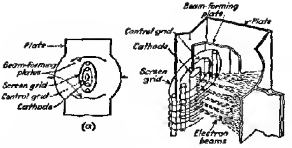


FIG 3 53 A horizontal cross section (a) and an artist's view (b) of a beam power tube

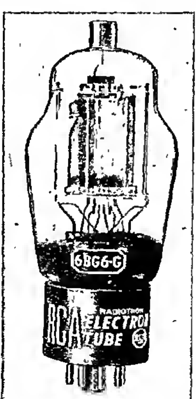


FIG. 3-54 Beam power tube. (*RCA.*)

serve to prevent any stray secondary-emission electrons from reaching the screen grid from the ends of the tube. The construction of the tube is further illustrated by the cross-sectional view of Fig. 3-53a and the photograph of Fig. 3-54.

Potential Distribution Curves. A potential distribution curve for a path intersecting a wire of the screen grid is shown in Fig. 3-55. A marked dip in the curve is observable between screen grid and anode, indicating the presence of a virtual cathode which prevents secondary-emission electrons from reaching the screen. For lower plate voltages the plate and virtual cathode act as a diode with plate current limited by space charge, giving the steep portion of the characteristic curves (Fig. 3-56) such as the region

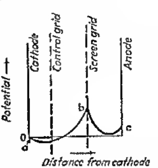


FIG. 3-55. Potential distribution curve of a beam power tube.

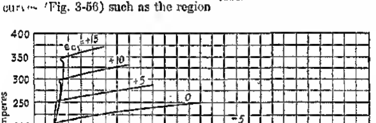


FIG. 3-56. Static characteristic curves of a beam power tube. These should be compared with the curves of a pentode, Fig. 3-48.

below $e_b = 50$ volts for a grid voltage of zero. For higher potentials the plate current is limited by saturation of the virtual cathode and is essentially independent of the plate voltage. This is illustrated by the flat portions of the characteristic curves (Fig. 3-56) lying above about $e_b = 50$ volts, for $e_c = 0$. Since the virtual

cathode has a comparatively large surface and since the design of the tube is such that the electrons flow in beams, the distance between the virtual cathode and the anode is nearly the same at all points. Thus conditions in the tube closely approximate those of the infinite parallel-plane electrodes assumed in plotting the ideal diode characteristic curves of Fig. 3-8, and the similarity between the beam-tube characteristic and those of the ideal diode is at once apparent. It is, of course, evident that the number of electrons reaching the virtual cathode and, therefore, the magnitude of the plate current under virtual cathode saturation conditions are under control of both control- and screen-grid potentials. Thus the family of curves of Fig. 3-56 is for various control-grid potentials rather than cathode temperatures as in Fig. 3-8. (The screen-grid potential was held constant.)

Comparison of the curves of Fig. 3-56 with those of the pentode (Fig. 3-48) shows the two to be similar except that those of the beam tube bend more abruptly at low plate voltages, thus extending the straight portions of the curves to even lower potentials. This permits a greater power output without distortion than is possible with the pentode.

Multigrid Tubes as Amplifiers. Multigrid tubes are commonly used as amplifiers in the same manner as triodes. The general circuit for a pentode amplifier is shown in Fig. 3-57 but details of performance will not be presented until Chap. 9. Beam tubes are used in a similar circuit but with the suppressor grid omitted. The resistor R allows the same source of direct current to be used for both plate and screen grid but with a lower voltage on the latter. The d-c source is usually a rectifier (see Chap. 7) rather than a battery as shown, and the use of a common source for both plate and screen is highly desirable. The condenser C maintains the screen grid at the same alternating potential as the cathode, a necessary condition as pointed out on page 71.

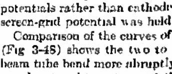


FIG. 3-57. General circuit for a pentode amplifier. The circuit for a beam tube is the same but with the suppressor grid omitted.

Comparison of Fig. 3-57 with Fig. 3-22 for a triode is of interest. Resistance R_L in the latter figure corresponds to the block marked "load" in Fig. 3-57. Alternatively this load might be inductive or capacitive as in Fig. 3-27, instead of resistive.

A useful viewpoint concerning the performance of multigrid tubes as amplifiers may be gained by first noting that the alternating voltage across the load is developed between plate and cathode in both triode and pentode (or beam tube) amplifiers and is therefore equal to the plate voltage. Thus a large alternating plate voltage must be developed at high outputs, and this voltage, added to the direct voltage, will result in a low instantaneous plate voltage at times t_1, t_2, etc., Fig. 3-58. As will be more fully shown in Chap. 9 this low voltage cannot be obtained in a triode amplifier without driving the grid positive, an undesirable procedure because of the resulting grid-current flow. Briefly, the grid must be driven positive because the plate in a triode is called upon to serve in the dual capacity of output electrode and as the source of attraction for bringing electrons over from the cathode through the control grid. As just stated, effective use of the plate as an output electrode requires that its potential be low at a certain time in each cycle (it should approach zero for maximum output), whereas its potential should remain reasonably high at all times if it is to attract electrons adequately. These are conflicting requirements, and a compromise must be effected whereby the plate voltage never drops too low at any point in the cycle. This reduces the maximum output of the amplifier but permits an adequate plate-current flow without driving the grid positive.

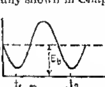


FIG. 3-58. Plate voltage curve for high output.

In pentode and beam tubes the plate serves as an output electrode only, while the screen grid attracts the electrons from the cathode. Thus the plate voltage may drop to very low values without affecting the current flow through the tube, permitting a much higher output voltage and resulting in an appreciably higher plate efficiency than with a triode. (Details of amplifier performance for both triode and pentode are presented in Chap. 9, beginning on page 328.)

4. MISCELLANEOUS HIGH-VACUUM TUBES[1]

There are a large number of other types of tubes, some of which differ only in the number or location of the electrodes, whereas others are of an entirely different type of construction from those

[1] This part may be omitted by those not interested in communication work.

already described in this chapter. A few of the more common are listed in the following sections with a brief note as to their construction and general characteristics

Pentagrid Converters and Mixers. Pentagrid converters are essentially tetrodes with two additional grids. The screen grid, however, consists of two grids tied together, G_3 and G_5 (Fig. 3-59), one on either side of the control grid G_4. This tube is used primarily in superheterodyne radio receivers (page 583) for combining the local oscillator and incoming radio signals, grids 1 and 2 being used in the local oscillator circuit, and grid 4 introducing the incoming signal from the antenna. Grid 2 may not be a grid at all but merely two rods to serve as the anode of the oscillator section.

The mixer tube of Fig. 3-60 is a variation of the pentagrid converter in which the processes of oscillation and mixing are entirely

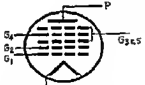


FIG. 3-59. Arrangement of electrodes in a pentagrid converter

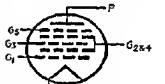


FIG. 3-60. Arrangement of electrodes in a pentagrid mixer tube.

separate (page 584). This tube is essentially a pentode in which the incoming signal is applied to the control grid G_1 but an additional control grid G_3 is provided to supply the local oscillator signal generated by a separate tube (instead of being generated in the same tube, as in the pentagrid converter). The screen consists of the two grids G_2 and G_4, tied together inside the tube, the additional screen G_2 being provided to shield G_1 from G_3. Grid 5 is the suppressor and is normally tied to the cathode through a connection external to the tube.

Duplex Tubes. A number of duplex tubes on the market are really two separate tubes in a single envelope, generally using different portions of a single cathode. An example is the duo-diode-triode, a triode tube with two small anodes mounted below the triode section. The diode section is normally used as a detector of radio signals; the triode section serves to amplify the a-f output

from the diodes. Often one anode of the diode section is used to supply automatic volume control, and the other demodulates the signal (see page 587). The same principle has been applied to a duo-diode-pentode.

Another tube built on this same principle contains a triode section and a pentode section. Both operate independently, the tube performing the same functions as a separate triode and pentode. Still another incorporates two triodes in a single envelope.

Space-charge-grid Tubes. A space-charge-grid tube is a tetrode in which the inner grid (closest to the cathode) is held at a slightly positive potential with respect to the cathode and the outer grid serves as the control element. When connected in this manner, the positive potential on the inner grid has the effect of neutralizing the space charge immediately surrounding the cathode. A large cloud of electrons then collects near the outer grid, forming a virtual cathode which is much closer to the latter grid than any physical cathode that could be constructed without danger of a short circuit between electrodes. The close proximity of the virtual cathode makes possible a much higher mutual conductance than in a conventional triode. Unfortunately the characteristic curves are less straight than are those of a conventional triode, and the current drawn by the space-charge grid tends to be rather excessive. This type of tube is, therefore, used only for special applications where its characteristics are peculiarly appropriate.

Acorn Tubes.[1] The operation of conventional tubes is very unsatisfactory at frequencies of the order of 100 Mc and over. This is due in large part to the capacitances between electrodes which provide a low impedance path in shunt with any external circuit and to the transit time of the electron which becomes appreciable in comparison to the time of a cycle.

An obvious solution to these problems is a reduction of tube dimensions. If all dimensions are reduced in proportion, the tube constants will be unchanged but the internal capacitances will be reduced and the transit time shortened. Two such tubes, known as *acorn* tubes, are shown in Fig. 3-61, (*a*) being a triode and (*b*) a pentode. They are 1½ in. from tip to tip and ¾ in. in diameter at the electrode-supporting ring (not including the protruding

[1] B. J. Thompson and G. M. Rose, Jr., Vacuum Tubes of Small Dimensions for Use at Extremely High Frequencies, *Proc. IRE*, **21**, p. 1707, December, 1933.

electrode wires) The electrodes are extremely small and supported by the lead wires, which are in turn supported in the glass ring with comparatively large spacing between wires. This construction keeps the interelectrode capacitance (including that of the lead wires) to a minimum. The grid-plate capacitance of the triode is 1.4 $\mu\mu f$ compared to 3.2 for the more conventional 56; the grid-cathode and plate-cathode capacitances are 1.0 and 0.6 $\mu\mu f$, respectively, compared to 3.2 and 2.2 for the 56.

The chief obstacle to the more extensive use of such small tubes is their inability to dissipate large amounts of energy from such small electrodes. Thus they are limited in their application to comparatively low-power installations.

FIG. 3-61a

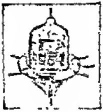


FIG. 3-61b

FIG. 3-61. Acorn type tubes, designed especially for frequencies of 100 to 600 Mc. (a) is a triode and (b) is a pentode. The over-all height of these tubes is but $1\frac{1}{2}$ in.

Lighthouse Tubes. Another approach to the problem of reducing the effect of the tube capacitances is to build the electrodes so that they become extensions of hollow-tube transmission lines, which in turn provide the desired external circuit elements. A picture of such a tube is shown in Fig. 3-62 where the plate terminal is the cap, the center ring is the grid terminal, and the cathode is connected to the base. Concentric conducting tubes are slipped over each of these terminals, and the transmission lines formed by these tubes constitute the external circuit (see Fig. 11-26). A brief description of an oscillator circuit using a lighthouse tube is given on page 477, the circuit being capable of oscillating at frequencies up to 1500 Mc.

Ultra-high-frequency Tubes. Many new tubes have been developed, especially during the Second World War, for use at frequencies from approximately 100 to 30,000 Mc (and even higher)

Most of these tubes utilize different principles of operation than those described in this chapter, largely because at these very high frequencies the transit time of the electron is no longer even approximately negligible as compared to the time of one cycle of the alternating voltage. Two outstanding examples are the velocity-modulated tubes, such as the klystron, and the various types of magnetrons. Since the ultra-high-frequency field is so extensive as to constitute a complete study in itself, no attempt will be made to describe such tubes in this book.

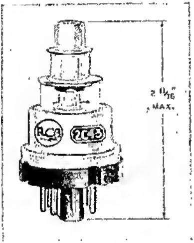


FIG. 3-62 Lighthouse tube for frequencies up to 1500 Mc.

Problems

3-1. The plate current in a certain diode is 10 ma at a plate voltage of 100. (a) If the tube is assumed to have infinite parallel-plane electrodes, what will be the plate current at 200 volts; (b) at 300 volts?

3-2. Assume that another diode with infinite parallel-plane electrodes is available with exactly the same constants as that in Prob. 3-1 except that the spacing between cathode and plate is only 90 per cent of the spacing in the tube of Prob. 3-1. What is the plate current at a potential of 200 volts?

3-3. The plate current in a certain triode is 30 ma when $E_b = 250$ volts and $E_c = -50$ volts. If μ is 3.5 and is assumed to be reasonably independent of electrode voltages, what plate current would you expect at $E_b = 300$ and $E_c = -40$ volts? (Assume that the plate current follows Child's law.)

3-4. The following modified form of Eq. (3-2) may be written for a triode with finite electrodes

$$i_b = k(e_c + e_b/\mu)^n$$

Plot a curve of i_b vs. $(e_c + e_b/\mu)$ on log-log graph paper and solve for k and n for the tube of Figs. 3-15 and 3-16 (μ was shown to be 15.4, page 57.)

3-5. Repeat Prob. 3-4 for the tube of Fig. 3-21. (μ was shown to be 3.6, page 58.)

3-6. Determine r_p for the tube to which the curves of Fig. 3-16 apply, for $e_c = -1$ volt, by measuring the slope of the curve. Take measurements at several values of e_b, and plot a curve of r_p vs. e_b.

3-7. Following the method of Prob. 3-6 determine g_m from Fig. 3-15, for $e_b = 50$ volts, and plot a curve of g_m vs. e_c.

3-8. A certain triode tube is to be operated at $E_b = 100$ volts, $E_c = -1$ volts, $I_b = 2.5$ ma. The tube coefficients at these voltages are $\mu = 20$, $r_p = 13{,}330$ ohms. A resistance load of 100,000 ohms is used in the plate circuit, and an alternating potential of $E_g = 0.1$ volt is impressed on the grid. (a) Compute I_p, E_p, E_{bb}. (b) Draw the vector diagram to scale.

3-9. A certain triode tube is to operate at $E_b = 250$ volts, $E_c = -8$ volts, $I_b = 8$ ma. The coefficients at these voltages are $\mu = 20$, $r_p = 10{,}000$ ohms, $g_m = 2000$ μmhos. An impedance of $4000 + j\,100{,}000$ ohms is inserted in the plate circuit, and an alternating grid potential of $E_g = 1$ volt is impressed. (a) Compute I_p, E_p, E_{bb}. (b) Draw the vector diagram to scale.

3-10. Repeat Prob. 3-9 assuming that the frequency is reduced to 10 per cent of the frequency to which the constants of Prob. 3-9 apply.

3-11. Repeat Prob. 3-6, for the tetrode curve of Fig. 3-40, for $e_{c1} = -3$ volts.

3-12. Repeat Prob. 3-7 for the tetrode curve of Fig. 3-39 for $e_{c2} = 120$ volts. Plot g_m vs. e_{c1}.

3-13. Repeat Prob. 3-6 for the pentode curve of Fig. 3-48 for $e_c = -10$ volts.

3-14. Repeat Prob. 3-8 for a pentode tube in which $r_p = 1.5$ megohms, $g_m = 1500$ μmhos, $E_b = 100$ volts, $E_{c2} = 75$ volts, $E_{c1} = -1.5$ volts, $I_b = 2.0$ ma, $I_{c2} = 0.5$ ma. (Compare results with those for the triode of Prob. 3-8. Note that both tubes have the same g_m.)

CHAPTER 4

GAS-FILLED TUBES

The presence of gas in a vacuum tube markedly alters its characteristics. This was pointed out rather briefly on page 48 where it was stated that gas tends (1) to increase the current flowing with a given applied potential (or, conversely, permit a lower applied voltage to produce a given current), (2) to cause damage to the cathode by positive-ion bombardment, and (3) to decrease the negative voltage at which the tube will break down and conduct current in the opposite (or undesired) direction. Generally speaking, the first of these effects is highly desirable, since it will decrease the amount of energy to be dissipated on the plate of the tube for a given load. Elimination, or at least reduction, of the undesirable features of the other two effects has made possible the development of gas-filled tubes which show marked superiority to high-vacuum tubes in many fields, notably in rectification and industrial control.

Two different types of discharges are to be found in gas-filled tubes: *glow discharges* and *arc discharges*, although the distinction between them is not great.[1] Glow discharges occur more commonly in tubes with cold cathodes; hot-cathode tubes are usually characterized by arc discharges.

Current Flow through Gases. If two cold electrodes are inserted in a vessel containing gas and a potential is applied between them, it might be expected that no current would flow since gases are normally considered to be good insulators. Actually, how-

[1] For a more complete analysis than can be given within the limited scope of this text see L. R. Koller, "The Physics of Electron Tubes," 2d ed., McGraw-Hill Book Company, Inc., New York, 1937; Herbert J. Reich, "Theory and Applications of Electron Tubes," McGraw-Hill Book Company, Inc., New York, 1944; William G. Dow, "Fundamentals of Engineering Electronics," John Wiley & Sons, Inc., New York, 1937; and John D. Ryder, "Electronic Engineering Principles," Prentice-Hall, Inc., New York, 1947.

ever, there are always a certain number of ions and free electrons present in the gas due to the action of cosmic rays, ultraviolet radiation, or other sources of excitation. These ions and electrons will flow toward the electrode of opposite polarity when a potential difference is applied and produce a current flow through the external circuit, Fig 4-1 At atmospheric pressures and at reasonable voltages this current is of the order of only a few microamperes and may be neglected for most engineering purposes; thus gases are usually treated as insulators If the potential difference is sufficiently increased, the gas will break down and cause a spark to jump between electrodes, followed by an arc if external circuit conditions permit.

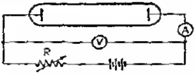


FIG 4-1 Circuit for obtaining characteristics of cold-cathode, gas-filled tube

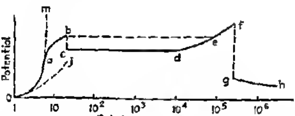


FIG 4-2 Voltage current characteristic of a typical cold-cathode, gas-filled tube

Reduction of the gas pressure in the vessel will permit an entirely different sort of phenomenon to take place when the potential is increased. At low potentials the current will increase in much the same manner as in a thermionic, high-vacuum tube; *i.e.*, it will first rise with increasing potential and will then level off as saturation is approached. This is the region *o* to *a*, Fig. 4-2. (Note that the abscissa is current rather than potential, as in Fig. 3-3, for example.)

As the potential is increased beyond point *a*, some of the electrons present in the gas reach velocities sufficiently high to ionize additional gas atoms by collision, resulting in what is known as the *Townsend discharge*. Many of the newly liberated ions and electrons also flow to the electrodes, thus increasing the current flow through the tube. Collision of an electron with a gas atom sometimes merely raises one or more of the orbital electrons into an orbit of higher energy and so leaves the atom in an *excited state* (see also page 6). Usually the excess energy contained in such an atom is immediately radiated in the form of visible and ultraviolet light (although the Townsend discharge is usually invisible) and the electron returns to its normal orbit. In some cases the energy may be lost only by contact with a surface where it may be transferred to another electron and so produce additional emission or may merely raise the temperature of the walls or electrodes of the tube. Such atoms are said to be in a *metastable state*.

Townsend discharges do not occur at atmospheric pressures since the mean free path of the electron (page 49) is then too short for the electron to attain sufficient velocity to ionize the gas. It is for this reason that reduced gas pressures must be used to produce the phenomena being described. Too great a reduction in gas pressure, on the other hand, will increase the length of the mean free path until very few electrons can collide with a gas atom in passing over to an electrode and the characteristic will be that of a high-vacuum tube, following a curve such as *oam*, Fig. 4-2.

The Townsend discharge is nonself-sustaining. If the ionizing source could be removed, the current flow would cease entirely, whereas if the ionization of the gas is increased, as by exposing the tube to ultraviolet light, the saturation current becomes greater, as shown by the dashed curve *j*. The curves *oj* and *oam* correspond to the curves 2 and 3 in Fig. 3-3 (except for the interchange of abscissa and ordinate). In Fig. 3-3 the saturation current is determined by the number of electrons liberated from a hot cathode, whereas in Fig. 4-2 saturation is determined by the number of electrons liberated by ionization due to cosmic rays or other sources.

Further increase in potential beyond point *b* may result in a sudden jump in current from *b* to *c*. Usually the external circuit contains resistance, Fig. 4-1, which prevents appreciable rise in current and, if an attempt is made to increase the potential beyond *b* by reducing this resistance, a slight rise in current will result and

the potential will fall from b to c. If the resistance is further reduced, the current will continue to rise, but no further increase in potential across the tube is required, even for very large increases in current (Note that the current in Fig. 4-2 is plotted to a logarithmic scale) This constant potential phenomenon is made use of in voltage regulator tubes (page 225).

The potential at b is sufficiently high to cause cumulative ionization, *i e.*, the electrons produced near the cathode liberate other electrons, as they travel toward the positive electrode, in such numbers that the process becomes cumulative and the discharge is self-sustaining and does not require an external ionizing source as in the Townsend discharge The exact process by which these new products of ionization produce additional electrons at or near the cathode is not fully understood but may include positive-ion bombardment, release of energy from metastable atoms at the cathode, and other means If the potential is maintained constant as the current exceeds b, this process continues out to the point e; but if the current is limited as in Fig. 4-1, the potential will drop until equilibrium exists in the tube as described in more detail in a later paragraph.

The potential at b is known as the *breakdown potential* of the tube. It is considerably higher than the ionizing potential of the gas where the *ionizing potential* may be defined as "the potential difference through which an electron must fall to attain sufficient energy to ionize an atom upon collision." The ionizing potential of most gases is of the order of 10 to 25 volts, whereas the breakdown potential may be several hundred volts.

The region *cdef* is characterized by a visible glow in the tube and on portions of the cathode. As the current is increased, the area of the cathode covered by this glow increases up to point d, the discharge being known as the *normal glow*. Beyond point d the cathode area is fully covered and the tube drop must increase to maintain the discharge, producing what is known as the *abnormal glow* Throughout the region *cdef* many excited and metastable atoms are produced and, as the displaced electrons drop back into their normal orbits, their excess energy may be released in the form of visible and ultraviolet light. Additional light is produced by the recombination of electrons and positive ions, usually at the walls and electrodes.

The amount of energy given up by an electron in dropping back

from one orbit to another is always the same for any given gas and is equal to one photon. This was shown (page 19) to be equal to $h\nu$ where ν is the frequency of the radiated wave. Thus each gas gives out a characteristic color.

At point f the character of the discharge again changes, in a manner not fully understood. The current density at the cathode increases suddenly, and the current tends to concentrate at one or more spots. Under these conditions an intense cloud of positive ions is formed around the cathode, and the electrons are able to flow from cathode to anode with an impressed potential which is only slightly higher (or even slightly lower) than the ionizing potential of the gas. The discharge in this region is known as an *arc discharge* and is characterized by very intense light, low tube drop, and high current density at the cathode surface. This is the type of discharge to be found in mercury-arc rectifiers and, in a modified form, in hot-cathode, gas-filled tubes. It will be more fully described in later sections of this chapter.

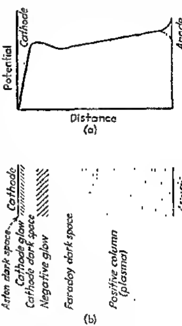


FIG. 4-3. Potential distribution and characteristic regions of a glow discharge.

Glow Discharges. The glow discharge found in the region $cdef$, Fig. 4-2, is of special interest in cold-cathode tubes which will be described in a later section of this chapter. While the discharge varies slightly in appearance as the current is increased from c to f, the general nature of the discharge remains the same. However the discharge is not uniform throughout the length of the tube but is divided into distinct regions, as shown in Fig. 4-3 for a pressure of the order of 1 mm of mercury, where (*a*) shows the potential distribution along the path of the discharge and (*b*) shows the light pattern drawn to the same dis-

tance scale. In Fig. 4-3*b* the positive column is seen to constitute the major portion of the normal discharge. This region is also known as the *plasma* and is characterized by a very low gradient and by the emission of the characteristic light of the gas used in the tube. It contains a large number of free electrons and positive ions in nearly equal densities. It is therefore very similar in characteristics to a good metallic conductor.

Next to the positive column is a region of low light intensity known as the *Faraday dark space*. This is followed by the *negative glow region*, the *cathode dark space* (or Crookes dark space), the *cathode glow*, and the *Aston dark space*.

It is through the cathode dark space that the major part of the drop in the tube occurs. This drop depends largely on the kind of gas used and to a lesser extent on the material used for the cathode. In glow discharges it ranges from about 75 to 400 volts.

The region of high gradient at the cathode is commonly known as a *positive-ion sheath*. In this region both positive ions and negative electrons are rapidly accelerated, the former toward the cathode and the latter toward the plasma. The electrons, being of smaller mass, are removed from the sheath more rapidly than are the ions; thus the term "positive-ion sheath" indicates the relative absence of electrons and not an increase in positive-ion density.

At the anode another region of higher gradient exists known as the *anode fall space*. The potential drop through this region is seldom more than a few volts and may be negative (as indicated by the dotted line).

The relative sizes of the various regions shown in Fig. 4-3 vary with pressure. In general a reduction in pressure is roughly equivalent to moving the anode closer to the cathode so that the positive column becomes shorter and eventually may disappear at low pressure. If the pressure is increased, the cathode dark space will disappear, the negative glow and cathode glow combining. As the pressure is further increased toward atmospheric, the discharge will eventually take the form of a series of sparks.

Arc Discharges. The arc discharge differs from the glow discharge largely in the magnitude of the drop through the cathode dark space, which is seldom much in excess of the ionization potential of the gas, *i.e.*, it seldom exceeds 25 volts and may be as low as 10 volts. A common method of securing this type of dis-

charge is to heat the cathode until it emits electrons thermionically in such numbers as to produce copious ionization (by collision) in the region immediately surrounding the cathode. The positive ions produced by this intense ionization set up a positive space charge which in turn permits the emitted electrons to pass freely away from the cathode and so constitute the arc current. Owing to the high ion density a positive-ion sheath may be formed around the anode as well as around the cathode, and arc discharges are commonly characterized by a negative gradient in the anode fall space.

Arc Back. As pointed out in the opening paragraph of this chapter there are two distinct limitations to the use of gas in vacuum tubes, of which one is the increased probability that a hot-cathode tube will conduct in the reverse direction (or arc back) when a high negative (or inverse) potential is applied to the anode.

However, it has been found that the negative voltage that a gas-filled tube is capable of withstanding is an inverse function of the gas pressure.[1] This is clearly shown by the curve of Fig. 4-4 where the inverse voltage rises from only 1500 volts at a pressure of 10 mm to over 30,000 volts at 0.03 mm pressure. In most practical cases the only limits to the voltage that may be applied by decreasing the pressure are imposed by the problems of insulation, as in the high-vacuum tube (which, of course, represents the practical limit of decreased pressure).

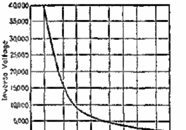


FIG. 4-4. Curve of inverse peak voltage as a function of mercury-vapor pressure.

Cathode Disintegration by Positive-ion Bombardment. The other limitation to the use of gas-filled tubes was given as destruction of the cathode by positive-ion bombardment. Positive ions formed by electron collisions with the gas atoms have a very much larger mass than the electron and, when attracted to the cathode, may strike with sufficient violence to cause serious damage to, or

[1] A. W. Hull, Gas-filled Thermionic Tubes, *Trans. AIEE*, 47, p. 753, March, 1928.

actual destruction of, the emitting surface, especially if high-efficiency composite surfaces are used. As the voltage to be rectified is increased, this problem of cathode disintegration by positive-ion bombardment becomes increasingly serious.

Cold cathodes and pure tungsten hot cathodes are not affected by this bombardment, but the high-efficiency, oxide-coated cathodes are quickly destroyed unless properly protected. The use of oxide-coated cathodes was therefore impractical in gas-filled tubes until Dr. A. W. Hull discovered, in 1923, that cathode disintegration could be almost entirely eliminated by keeping the tube drop below a certain critical value.[1] This value was found to vary from about 20 to 25 volts for most of the inert gases, which is the potential at which double ionization occurs. As long as the gas is singly ionized (only one electron removed from the atom), the positive ions do not strike with sufficient force to injure the cathode surface seriously. At the critical potential referred to, two electrons are removed from the atom which is then attracted to the cathode with twice the force owing to its double charge. The increased velocity of impact of the positive ions, as the point of double ionization is passed, is sufficient to disintegrate an oxide-coated cathode surface rapidly. Thus either the drop must be kept below about 22 volts, or the less efficient tungsten cathodes must be used.

1. DIODES

Hot-cathode, Mercury-vapor Tubes. The studies of arc-back and cathode disintegration referred to in the opening sections of this chapter led to the development of a satisfactory gas-filled tube containing liquid mercury and mercury vapor in equilibrium at a pressure of 1 to 30 microns. This tube uses a hot cathode to produce emission and is therefore known as a *hot-cathode, mercury-vapor tube*, which will be frequently abbreviated in this book to simply *mercury-vapor tube*. Current flow through this tube is in the nature of an arc discharge which gives off the characteristic blue glow of mercury.

The inverse voltages that the tube will withstand are reasonably high and yet the internal drop is but 15 volts. Oxide-coated

[1] A. W. Hull and W. F. Winter, The Volt-ampere Characteristics of Electron Tubes with Thoriated Tungsten Filaments Containing Low-pressure Inert Gas, *Phys. Rev.*, **21**, p. 211, 1923 (abstract).

cathodes and heat shielding (see page 104) may be used satisfactorily with consequent increase in emission efficiency. As previously stated, high-voltage, high-vacuum tubes are built almost entirely with tungsten filaments because oxide-coated or thoriated cathodes are too likely to be spoilt by positive-ion bombardment from the small amount of gas that may be present. In the mercury-vapor tube the drop is kept below 22 volts (the potential at which cathode disintegration occurs due to double ionization of the mercury ions) so that the more efficient oxide-coated emitting surface may be used.

The anode of the mercury-vapor tube usually consists of a small target; the cathode may be a heavy coiled or zigzag ribbon or a heat-shielded cylinder (see page 104), although the very small sizes more nearly resemble the high-vacuum tubes of equivalent rating. All but the larger sizes are built in glass containers, the larger ones using the metal anode as the envelope. The tube drop is so low that water cooling is not normally necessary as in the case of the larger sized high-vacuum tubes.

Mechanism of Current Flow. Application of a positive potential to the plate of a mercury-vapor tube will cause a flow of electrons from cathode to anode just as in the high-vacuum tube. These electrons cannot travel unimpeded as in the high-vacuum tube, however, but will collide with the gas atoms that occupy the space between electrodes. If the potential across the tube exceeds the ionizing potential of the gas, some of these collisions will be sufficiently violent to produce ionization, and a discharge will take place in the tube.

The space charge between the cathode and the anode of a gas-filled tube, unlike that of a high-vacuum tube, is produced by both negative electrons and positive ions. Furthermore, the positive ions, being many times heavier than the electrons, travel at a much lower velocity. Therefore, although they are too few in number to add appreciably to the actual flow of current, there are approximately as many of them in the space within the tube at any instant as there are of electrons, and the resulting space charge is practically zero except near the electrodes and walls.

The potential distribution curve, shown in Fig. 4-5, is very similar to that of the glow discharge, Fig. 4-3. Most of the drop occurs near the cathode, in the region *a*, with a small drop at the anode. The gradient throughout the plasma is very small owing

to the nearly equal concentration of electrons and ions The potential of the plasma is only 10 to 15 volts higher than that of the cathode, in contrast to the glow discharge where it may be as high as several hundred volts

Since the gradient near the cathode is positive, as shown in Fig. 4-5, it might be expected that all electrons emitted from the cathode would be drawn toward the plate, thus making the plate

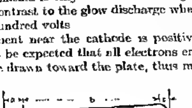


FIG 4-5 Potential distribution curve of mercury-vapor diode.

current equal to the emission current of the cathode. Actually the current flowing in the circuit of Fig 4-6 is nearly equal to the applied potential divided by the load resistance, since the tube drop is very small, and any change in load resistance immediately alters the current demanded by the circuit. This is accompanied by small changes in gradients within the tube such that the flow of

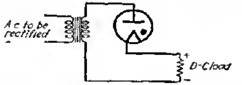


FIG 4 6 Illustrating the use of a mercury-vapor tube as a rectifier

electrons to the anode is maintained just equal to the demand. Because of this characteristic, gas-filled tubes should never be operated without sufficient resistance in series to limit the current to a safe value.

The anode drop in Fig. 4-5 is negative, producing a positive-ion sheath which is characteristic of mercury-vapor tubes. The drop through this sheath is sufficient to prevent most of the electrons from leaving the plasma and thus maintains the balance between

ions and electrons which is characteristic of the plasma. The potential will decrease somewhat as the load current is increased to permit a larger number of electrons to be drawn off and will even become positive at high currents.

Normally, changes in load have but very little effect on the plate voltage (*i.e.*, tube drop). Evidently it is primarily the plasma potential (region *b* of Fig. 4-5) that changes a small amount to produce the variations in anode drop made necessary by changes in load. The plasma potential, therefore, assumes a point of equilibrium that just balances the inflow of electrons from the cathode, the formation and recombination of ions within the tube, and the outflow of electrons to the anode.

Maximum Permissible Current Flow. Decreasing the resistance in series with either a high-vacuum or a gas-filled tube will result in an increase in current, up to the point of cathode saturation. In the high-vacuum tube this will be accompanied by an increasing plate voltage (or tube drop), according to Child's law (page 44), whereas in the gas-filled tube the tube drop will decrease slightly. Any attempt to increase the current beyond the saturation point by still further lowering the external resistance will, in the case of a high-vacuum tube, be unsuccessful; the tube drop will continue to increase, but the current cannot exceed the emission of the cathode. If the same attempt is made with a gas-filled tube, on the other hand, it will be found that the current will *continue to increase*, which must indicate increased ionization, since the cathode emission is fixed by its temperature. Normally the number of electrons that ionization contributes to the total flow is very small compared with the contribution of the cathode; but if the load demand tends to exceed the cathode emission, the *anode voltage will rise sharply until the increased ionization supplies the demand.* At these higher voltages many atoms will be doubly ionized, resulting in destructive bombardment of the cathode. *Thus no attempt should be made to increase the plate current above the normal emission of the cathode.*

An excellent test for determining the condition of the cathode is based on the foregoing facts. The tube is placed in a test socket, and a direct current of 1½ to 2 times normal is passed through it. The tube drop is measured; and if in excess of about 22 volts (the potential at which double ionization of mercury vapor occurs), the tube has nearly reached the end of its useful life. Where con-

tinuity of service is an important factor, *e.g.*, as in radio broadcasting, this is a valuable feature.

Heat-shielded Cathode. The emission efficiency (ratio of emission current to heating power) of the hot-cathode, mercury-vapor tubes may be greatly increased by heat shielding the cathode. Dr. A. W. Hull[1] found that, if the electron-emitting surface of an indirectly heated cathode was coated on the *inner* surface (Fig. 4-7) toward the heat, instead of on the outside, the emission effi-

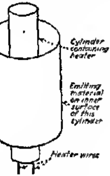


FIG. 4-7. Heat-shielded cathode showing the increased emitting surface available.

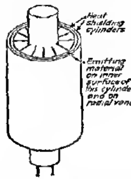


FIG. 4-8. Additional heat shielding may be obtained by placing shiny metal cylinders around the cathode. The emitting surface of this cathode has been still further increased by inserting radial fins, coated with oxide.

ciency was increased about 3:1. He also found that placing two shiny nickel rings around the cathode, as in Fig. 4-8, still further decreased the heat loss so that the loss was only one-sixth as much as with the ordinary outside-coated cathode.

A number of radial vanes are also shown in this figure. These vanes are coated with emitting material to provide an increased emitting surface. The heat loss is not affected by the presence of these vanes, so that a much greater emission is obtained with no increase in heating power. Hull states that the emission efficiency of such a cathode may be as much as twenty-five times as great as that of the ordinary outside-coated cylinder, such as that of Fig. 1-4. Heat shielding may be, and often is, applied to filamen-

[1] See Hull, *loc. cit.*

tary-type cathodes as well as to the indirectly heated type. The usual construction is to enclose the filament in a cylinder as in Fig. 4-9.

Heat shielding tends to increase the life of a cathode as well as its efficiency. Coated cathodes finally lose their usefulness as a result of the evaporation of the emitting surface. When shielding is used, an evaporating atom has much less chance of escaping far from the emitting surface and, since it is neutral and will not be attracted by the plate, it is far more likely to fall back into the surface than to be permanently evaporated.

It is impossible to use heat-shielded cathodes in high-vacuum tubes, as the electrons cannot escape readily from inside the shield and so will build up a very large space charge which the plate is unable to overcome. In mercury-vapor tubes the positive ions that are formed tend completely to neutralize the space charge throughout the tube, and the plate is therefore able to draw the electrons out from inside the shield.

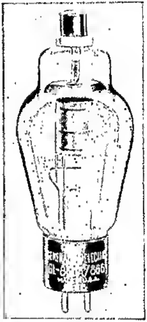


Fig. 4-9. Mercury-vapor diode rated at 1.0 amp. peak plate current and 10,000 volts inverse peak. (*General Electric Co.*)

Temperature Limits. It is important that the operating temperature of mercury-vapor tubes be kept within certain prescribed limits, too low a temperature being as injurious to them as is one too high. They are built with an excess of liquid mercury in the bulb to maintain a saturated condition of the vapor, and the temperature of this condensed mercury determines the pressure in the tube. If this pressure becomes too low, the negative space charge is not sufficiently well neutralized, and the tube drop becomes excessive, whereas an increase in vapor pressure will decrease the inverse voltage that a tube will withstand. Figure 4-10 shows curves taken on a typical mercury-vapor tube. The lower curve shows the relation between tube drop and condensed-mercury tempera-

ture, and the upper curve gives the flashback, or peak inverse voltage. A 15° difference in temperature is expected between the temperature of the condensed mercury and ambient. The tube-drop curve is practically the same for all hot-cathode, mercury-vapor tubes, being virtually independent of the amount of current flowing, as long as the current demand is large enough to cause ionization but not so large as to exceed the cathode emission. It may also be seen that the minimum temperature at which the condensed mercury may be operated, without causing cathode disintegration, is about 20°C, whereas the maximum temperature depends entirely on the inverse voltage of the circuit in which the tube is to be operated. Tubes are generally designed for a maximum normal operating temperature of 75°C. The temperature of the tube may usually be kept within this upper limit by allowing free circulation of air around the base of the tube; or if this is not possible, a blower may be provided.

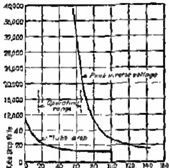


Fig. 4·10. Tube-drop and arc back voltages of a hot-cathode, mercury-vapor tube as a function of the mercury temperature.

Mercury-arc Tubes. Construction Details. The mercury arc is one of the oldest of the modern gas-filled tubes, having been used for battery charging as many as 30 years ago. In more recent years much development work has been done, and mercury-arc rectifiers have been constructed in sizes sufficiently large to supply direct current to streetcars and similar loads, capacities of several thousand kilowatts being quite common. A complete treatise on the mercury-arc rectifier is entirely beyond the scope of this book, but a brief discussion of the principles under which it operates will be given.[1]

The mercury-arc tube consists fundamentally of a cathode pool of mercury and one or more anodes, all enclosed in some sort of airtight, evacuated chamber. The smaller models are enclosed in a

[1] For a more complete discussion see D. C. Prince and F. B. Vogdes, "Mercury-arc Rectifiers and Their Circuits," McGraw-Hill Book Company, Inc., New York, 1927.

glass bulb (Fig. 4-11); the larger units are built in iron tanks (Fig. 4-12).[1] In Fig. 4-11 the mercury-pool cathode is located in the lower portion of the bulb, and the two anodes (full-wave, single-phase rectifier) are located in the long arms projecting to either side of the bulb. The large glass dome is the condensing chamber where the mercury vapor is condensed and allowed to drain back into the pool. The two small electrodes immediately above the pool are "keep-alive" anodes used to maintain the arc. The small tube of mercury extending down to the right side of the pool is the starting electrode by means of which the arc is struck.

The need for keep-alive electrodes is more fully covered in Chap. 7 (page 212). Briefly, however, the mercury-arc tube is ready for service only so long as the arc on the surface of the mercury pool is maintained, as it is this arc which causes emission of electrons from the cathode. If the load on the tube dropped to zero even for a fraction of a second, the arc would be extinguished and the tube would be out of service.[2] Thus a small arc is maintained between the keep-alive anodes and the cathode pool at all times to insure continuous operation of the tube. The power consumed by this arc is of the order of a few hundred watts at most, even in the largest tubes; thus it does not materially lower the efficiency of reasonably large rectifiers.

The metal-tank tube normally contains six or more anodes and is operated from a three-phase supply. The anodes are protected from arc back by shields and baffle plates, which both prevent splashing of the mercury on to the anode (and so perhaps forming a cathode spot) and increase the length of the arc path. Water cooling, to keep the mercury-vapor pressure within the proper operating limits, is provided by a water jacket or by cooling coils located in the condensing chamber or both.

The construction of a tube having several anodes and a single cathode in one container is peculiar to the mercury-arc rectifier. Such a construction would be impossible in a high-vacuum tube, since those anodes which were negative at any given instant would set up such a high negative space charge as to render the positive

[1] The glass tubes are obsolete today. Even the iron-tank type is being largely superseded by the ignitron (p. 112).

[2] It is necessary only that current flow through the tube cease for a length of time sufficient to permit deionization of the gas. This is usually not more than a few hundred microseconds.

anodes powerless to draw electrons. In the mercury-arc rectifier, on the other hand, the negative space charge is neutralized by the presence of positive ions.

As a matter of fact only in the mercury-arc tube is it at all desirable to use a single cathode with several anodes in a single envelope. Generally, unit construction is preferred because of the decreased probability of arc back and the greater ease of making replacements, thus where individual high-vacuum, mercury-vapor or ignitron tubes are used one tube may be replaced at any time without affecting the others. In the mercury arc, however, the necessity for maintaining a keep-alive circuit, with its consequent loss, makes the use of a single cathode highly desirable.

Mechanism of Current Flow. The mercury-arc tube is put into service by applying voltage to the keep-alive anodes and momentarily closing a circuit through the mercury pool, then breaking it to start the arc. In the glass tube of Fig. 4-11 the arc is struck by tipping the bulb until the mercury from the main pool flows over into the starting pool (seen projecting down from the lower right side of the bulb), permitting current to flow through an external circuit. The tube is then returned to its normal position, and the arc is formed as the pools separate. In the iron-tank tube various methods of striking the arc are in use. One such is to energize a solenoid on the top of the tank, inserting a starting electrode into the cathode pool. Deenergizing the solenoid allows a spring to withdraw the electrode, and the arc is struck.

A cathode spot is formed on the surface of the mercury pool as the arc is struck, and this spot is the source of electron emission. Mercury atoms are also evaporated and constitute a vapor or gas in the bulb. The electrons are attracted to a positive anode as soon as they are emitted, and in their passage through the tube they collide with and ionize some of these gas atoms.[1] The positive ions so formed are attracted to the cathode and neutralize the space charge as in a mercury-vapor tube.

The temperature of the cathode spot is apparently too low to produce thermionic emission and, although some secondary emission is undoubtedly produced by positive-ion bombardment, most

[1] If there is no load on the tube, there will be no positive anode except in the keep-alive circuit. One or the other keep-alive anode is positive at all times.

of the current appears to be due to high-field emission[1]. The positive ions are attracted to the cathode in such numbers that they set up a very high positive gradient, of the order of millions of volts, so that the electrons may be extracted from the liquid mercury by electrostatic forces. Since the total drop through the

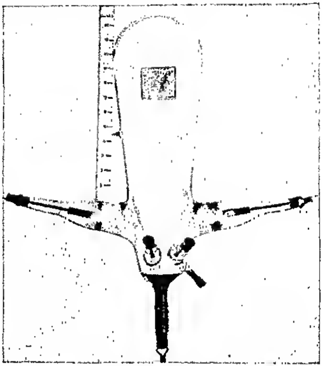


FIG. 4-11. Obsolete type of glass-bulb, mercury-arc rectifier tube, 50 amp d-c.

tube is only about 25 to 40 volts, this layer of positive ions must be in the nature of a very thin film on the surface of the mercury.

Arc Back. The arc-back voltage, as in the case of the mercury-vapor tube, is a function of vapor pressure which must therefore

[1] Dr. Irving Langmuir, Positive-ion Currents in the Positive Column of the Mercury Arc, *Gen. Elec. Rev.*, **27**, 1923.

be kept within certain limits. To do so requires adequate cooling, since the pressure is determined by the temperature of the coolest part of the bulb, in this case the condensing chamber. Glass-enclosed rectifiers are cooled by permitting free circulation of air around this chamber or by adding forced ventilation; iron-tank rectifiers are cooled by circulating water through suitably placed cooling coils. Obviously, the amount of cooling required increases with the load current. An increase in current causes an increase in the intensity of the cathode spot and, therefore, an increase in pressure, and unless sufficient cooling is provided, the pressure may soon rise to the point where arc back will occur. The maximum current that a mercury-arc tube will safely pass is determined solely by the amount of heat that may be dissipated, the number of electrons emitted increasing automatically with the demand by virtue of increased cathode-spot activity. In this it differs from the mercury-vapor tube wherein the current cannot be permitted to exceed the total emission of a fixed-temperature cathode, no matter how much cooling is provided.

As already explained, the arc in the mercury-arc tube must be maintained continuously while the tube is in service. Thus positive ions are present at all times; and when an anode is at a negative potential, it tends to attract these ions and thus pass current in the reverse direction. In commercial tubes this possibility is minimized by locating the anodes a comparatively long distance from the cathode and by placing shields and baffle plates between the cathode and each anode. The ions, being heavier, move much more slowly than the electrons and therefore do not have time to reach the anodes in any considerable quantity during the negative half cycles, but the lighter electrons reach them easily and in large numbers during the positive half cycles.

The drop in a mercury-arc tube tends to be higher than in a mercury-vapor tube. This is due in part to the greater spacing between cathode and anode but perhaps even more to the fact that the energy required to produce emission in the former tube is derived entirely from the cathode drop whereas in the latter a large part of the energy required for emission is produced by the cathode-heating source.

Mercury-arc vs. Mercury-vapor Tubes. From the foregoing discussion it is evident that the mercury-arc and mercury-vapor tubes have many points in common. Both have a very low inter-

nal drop which is constant regardless of the load; both will withstand a reasonably high inverse voltage without breaking down; and both require a close control of the vapor pressure for satisfactory operation. They have one very distinct difference, however:

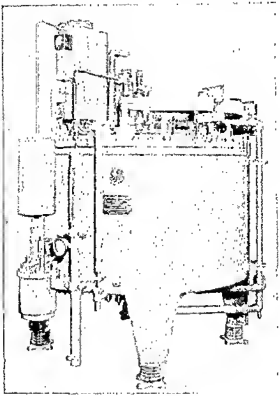


FIG. 4-12. Mercury-arc rectifier rated at 2500 kw, 3000 volts. (*General Electric Co.*)

the mercury-vapor tube is always ready for service as long as the cathode heater is energized regardless of how intermittent the load may be, whereas the mercury arc must be started by striking an arc that must be maintained either by a continuously flowing load current or by a special keep-alive circuit. On the other hand, this difference in their cathodes makes it possible to draw almost

unlimited current for short periods from the mercury-arc tube without causing injury, whereas the cathode of the mercury-vapor tube will be seriously damaged or destroyed in a short time if an attempt is made to draw more than its rated current. In other words, the maximum instantaneous current that may be passed through a mercury-vapor tube is determined by the emission obtainable from the hot cathode as well as by the temperature of the bulb, and the maximum current that may be passed by a mercury-arc tube is determined solely by the ability of the cooling system to keep the vapor pressure from becoming excessive.

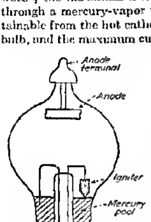


FIG. 4-13. Sketch showing principal parts of an ignitron.

Ignitron. Evidently a tube that will incorporate the ready-to-start characteristics of the mercury-vapor, hot-cathode tube with the heavy current-carrying capacity of the mercury-arc tube should prove of inestimable value. Such a tube is the ignitron. Fundamentally the ignitron is a mercury-arc tube consisting of a mercury pool in the bottom of an evacuated bulb with a suitable anode above. In addition it contains an igniting electrode which, with its associated circuit, serves to strike the arc at a given time as determined by external circuit conditions.

The operation of ignitrons is dependent on a principle discovered by Slepian and Ludwig.[1] The igniting electrode is made of a suitable refractory material and so mounted as to make contact with the mercury pool. A large current is passed between the igniter and the mercury, causing a small spark to strike which is carried over into an arc between anode and cathode if a suitable potential is applied to the anode at the time the spark is struck. Once the arc is established, the igniter circuit is opened to prevent waste of energy.[2]

[1] J. Slepian and L. R. Ludwig, A New Method of Starting an Arc, *Elec. Eng.*, 52, p. 605, September, 1933.

[2] Circuits for using the ignitron and for starting and stopping the igniter current are given in Chap. 7 beginning on p. 206.

The general construction features of an ignitron are shown in Fig. 4-13 where the igniter electrode may be seen in contact with the mercury pool. A close spacing between anode and cathode is made possible by the reduced danger of arc back resulting from the absence of a keep-alive arc. The close construction insures a very low tube drop and consequent higher efficiency. A commercial tube is shown in Fig. 4-14.

Fig. 4-14. Ignitron. (*General Electric Co.*)

The larger sized tubes are commonly enclosed in a metal tank with a water-cooling jacket, as in Figs. 4-15 and 4-16. The baffles and shields shown in Fig. 4-15 are inserted to decrease the probability of arc back. Unfortunately their presence also increases the tube drop, and in tubes intended for low-voltage applications they are omitted. An example of such service is in welding control where the impressed voltage is comparatively low, which means not only that the inverse voltage is low but that the tube drop must be kept to a minimum if it is not to be an appreciable percentage of the impressed voltage.

Figure 4-17 shows 12 ignitrons assembled into a complete rectifier rated at 3000 kw, 600 volts. (Circuits for this service are presented in Chap. 7.)

Comparison of Ignitron with Mercury-arc and Mercury-vapor Tubes. The principal fault of mercury-arc tubes was shown to be the need for a keep-alive circuit to maintain the arc, thus increasing the danger of arc back due to the continuous existence of positive ions in the tube chamber. The principal fault of mercury-vapor tubes was seen to be the limitation in permissible overload current imposed by the maximum emission of the cathode. The ignitron very successfully eliminates both these faults. The overload capacity is the same as that of the mercury-arc, and it is not necessary to sustain the arc, the tube being ready for service at all times.

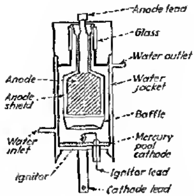


FIG. 4.15 Sketch showing general construction of a water-cooled ignitron.

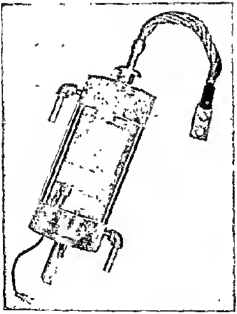


FIG. 4.16 Cutaway view of a water-cooled ignitron for use in welding service. Note the absence of baffles and shields (*Westinghouse Electric Corp.*)

Unfortunately these advantages of the ignitron are not obtained without cost, since it is necessary to use auxiliary tubes and circuits to energize the igniter (see page 206). As a result the ignitron is rapidly replacing the mercury arc in heavy-duty service where the cost of this auxiliary equipment is not too large a percentage of the total investment, but the mercury-vapor tube continues to be superior for lighter services.

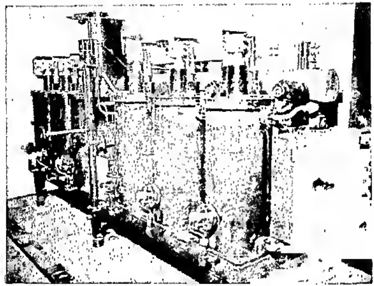


FIG. 4·17. Ignitron rectifier using 12 ignitrons, rated at 3000 kw, 600 volts. (*General Electric Co.*)

Tungar Tubes. The tungar tube (Fig. 4-18) was developed to supply a demand for a high-current, low-voltage tube to charge storage batteries or for other similar services. It is constructed with a rather heavy, coiled, thoriated-tungsten filament mounted horizontally between two vertical supports and an anode consisting of a small graphite button or target mounted above the filament. The bulb is well evacuated, and then argon gas is admitted to a pressure of about 5 cm of mercury. A ring of magnesium is secured around the plate, and after evacuation it is volatilized by heating of the anode. This magnesium *getter* serves the same purpose in this tube as in the high-vacuum tubes; it deposits on the

glass walls of the bulb and absorbs any active gas, such as oxygen, that may be present after evacuation or released from the metal parts of the tube during operation.

The filament of the tungar tube is operated at a much higher temperature than can be used in a thoroughly evacuated bulb so that its emission efficiency (milliamperes of emission per watt of filament heating power) is many times that of the filaments used in high-vacuum tubes. This high operating temperature would cause rapid evaporation, were it not for the presence of the gas. It has been found that cathode evaporation is practically negligible in the presence of an inert gas at the pressure used in this tube. Consequently a very large increase in emission is achieved in this tube without an excessive increase in heating energy.

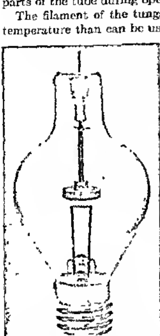


FIG. 4.18. Five ampere tungar tube (*General Electric Co.*)

Mechanism of Current Flow. The tungar tube conducts current in a manner very similar to that of the hot-cathode, mercury-vapor tube. The electrons that flow from cathode to anode when the latter is positive collide with gas atoms, forming positive ions which neutralize the space charge; thus saturation current may flow with a tube drop of the order of 10 volts. The tube drop may even be lower than the ionization potential of the gas once the tube has started to conduct. The cause of this very low drop is not entirely clear, but it is undoubtedly due to a combination of various ionization processes within the space between cathode and anode.[1]

The positive ions in the tungar, as in the mercury-vapor tube, do not constitute an appreciable portion of the current flow, since,

[1] A more detailed discussion of this phenomenon is given by L. R. Koller, "The Physics of Electron Tubes," 2d ed., p. 143, McGraw-Hill Book Company, Inc., New York, 1937.

because of their greater mass, their velocity is very much less than that of the electrons. As a consequence, virtually all the current through the tube is due to electronic emission from the cathode.[1]

During the half cycle that the plate is negative with respect to the filament, no electrons will be attracted to the plate, no ionization will occur, and the tube will present an open circuit to the flow of current. However the negative voltage applied to the plate must not be raised too high, or the gas will be ionized by the few free electrons always present, and conduction in the reverse direction will take place. This type of tube has not been found very satisfactory on voltages much over 100 volts.

Figure 4-19 shows the circuit commonly used with this tube for battery charging. The lower end of the secondary winding is wound with very heavy wire to supply the filament current which, for the 5-amp bulb, is about 14 amp, while the upper end supplies the plate voltage. The efficiency is low, since about 30 to 40 watts is required for filament heating and about 50 watts is lost inside the tube when delivering 5 amp to a 6-volt battery. The efficiency is, however, much better than when charging batteries from a 110-volt, d-c line through a resistance and compares rather favorably with that of a motor-generator set when the difference in first cost and maintenance is considered.

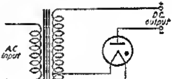


FIG. 4-19. Circuit of a tungar rectifier used for battery charging.

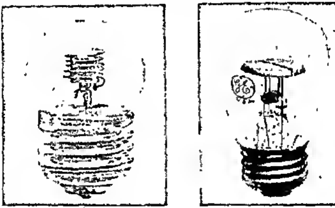
Cold-cathode Tubes. Tubes with cold cathodes may be operated in the region c to d, Fig. 4-2. The current-carrying capacity of such tubes is small, but they are characterized by a potential drop which is virtually independent of the current flowing. Thus they find extensive applications in voltage regulators (see page 225). Commercial tubes are usually made with tube drops of 75, 105, or 150 volts and with current-carrying capacities of from 5 to 50 ma. They are normally filled with inert gas such as argon, helium, or neon.[2]

[1] For a more detailed discussion of the mechanism of current flow through gas see the discussion on mercury-vapor tubes (p. 101).

[2] For a more detailed treatment of the performance of this type of tube, see George M. Fitzpatrick, Characteristics of Certain Voltage-regulator Tubes, *Proc. IRE*, 35, p. 485, May. 1947.

Since neither cathode nor anode is heated in a cold-cathode tube, either electrode may serve as the cathode, depending on the polarity of the applied emf. Cold-cathode tubes are, therefore, inherently bilateral conductors; *i.e.*, they will pass current equally in either direction. Examples of such tubes are the small neon lamps that are used as night lights or indicators of live circuits (Fig. 4-20).

It is possible to make cold-cathode tubes conduct more readily in one direction than in the other by constructing the electrodes



(a) (b)

FIG. 4-20. Small neon lamps. (a) 1 watt, the electrodes consisting of a cylinder and concentric spiral; (b) 2 watts, with identical electrodes. (*General Electric Co.*)

with different cross-sectional areas. For example, if one electrode consists of a small wire and the other is of relatively large cross section, the current flow will be greatest when the small electrode is positive. Apparently the tube drop is nearly independent of the magnitude of the current flowing as long as the current density on the surface of the cathode does not exceed a given amount.[1] However, if the current required by the load is such as to cause the current density to exceed this critical value on the small electrode but not on the large one, a much higher voltage would have to be applied in one direction than in the other to cause the same

[1] For further information see A. E. Shaw, Cold Cathode Rectification, *Proc. IRE*, 17, p. 849, May, 1929.

current to flow on both polarities of the supply. Since the voltage of both halves of the a-c supply is normally the same, the current flow in one direction is much greater than in the other and a net direct, or rectified, current results.

Cold-cathode tubes are commonly used as voltage regulators, as in the very elementary circuit of Fig. 4-21, where the voltage across the load will be maintained at the rated voltage of the tube provided the supply voltage is appreciably higher. If the load current increases, tending to reduce the voltage, the tube will immediately draw less current in order to satisfy the current-voltage characteristic curve of Fig. 4-2, so that the total current through the resistor R remains essentially constant regardless of the magnitude of the load current. A reduction in load is, of course, accompanied by an increase in the current flowing through the tube. On the other hand an increase in the supply voltage will tend to raise the output voltage, but the current through the tube will immediately rise until the increased drop through R again leaves the same output voltage across the tube and load. The degree of constancy of the output voltage is evidently dependent upon the flatness of the characteristic curve of Fig. 4-2 throughout the operating range c to d. In commercial tubes the drop varies only a few volts from minimum to maximum current. Regulator circuits may be designed using these tubes in conjunction with high-vacuum tubes which will maintain the output potential constant to within a fraction of a volt. Such circuits are described in Chap. 8.

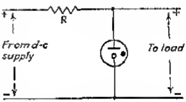


FIG. 4-21. Elementary voltage-regulator circuit using a cold-cathode tube.

2. TRIODES

The insertion of a grid into a gas-filled tube provides a means of controlling the flow of plate current just as in the case of a high-vacuum tube. Nevertheless the nature of this control is considerably different from that in a high-vacuum tube. In the latter the grid is able not only to stop and start the current but also to control its magnitude. In the gas-filled tube, on the other hand, the grid can only start the current; it can neither stop the flow of current once it has started nor alter its magnitude.

Thyratrons. A mercury-vapor, hot-cathode tube with grid control is known as a *thyratron*. The internal construction of a 5-amp tube of this type is shown in Fig. 4-22. The anode and cathode are of the same general type of construction as those of the mercury-vapor diode (page 101), although the cathode is of the indirectly heated type in nearly all thyratrons except those of very small size.

The grid of the tube shown in Fig. 4-22 consists of a metallic cylinder completely surrounding the cathode and anode, either perforated or solid. A baffle plate is inserted between cathode and anode and made an integral part of the grid. A small opening is provided in this baffle to permit the passage of electrons. The reason for this type of construction will be apparent shortly.

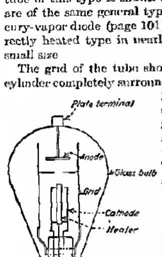


FIG. 4-22. Cross-sectional view of a small thyratron.

Application of a sufficient negative voltage to the grid, with the plate voltage off, will prevent any electrons from reaching the plate after the plate circuit is energized. If the grid voltage is then gradually made more positive (the plate circuit now being energized), a point will be reached where electron flow to the plate will suddenly begin. This value of grid potential is known as the *starting voltage* and is a function of the temperature of the mercury and of the plate potential, being more negative as the plate voltage is made more positive. Its absolute value may be either positive or negative, depending upon the tube construction.

Making the grid voltage equal to the starting value permits a flow of electrons to the plate which will cause ionization of the mercury vapor just as in the mercury-vapor diode. The positive ions thus formed will then neutralize the space charge and permit unimpeded passage of the electrons to the plate.

Suppose that a thyratron is placed in the circuit of Fig. 4-23 and the potential of the C battery at point 1 is sufficient to prevent the passage of electrons to the plate while point 2 is at a potential more positive than the starting voltage. If the grid switch is in position

1 at the time when plate voltage is applied, no current will flow; but as soon as it is thrown to position 2, conduction will begin and the relay will be energized. If an attempt is then made to stop the flow of current by throwing the switch back to position 1, it will be found that the grid has lost control of the plate current and that the only way in which the flow of current may be stopped is by opening the plate circuit itself.

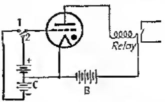


FIG. 4-23. Illustrating the impossibility of stopping the flow of current through the plate circuit of a thyratron by grid control alone.

Potential Distribution Curve. The potential distribution curve of a thyratron while conducting is illustrated in Fig. 4-24 for a negative grid potential. Figure 4-24*a* shows the distribution of potential along a line intersecting a portion of the metallic grid structure, whereas curve *b* represents conditions along a line drawn through an opening in the grid structure. A positive-ion sheath forms around the grid structure similar to the anode and cathode sheaths described at the beginning of this chapter (pages 98-102); thus the influence of the grid does not extend very far from its own surface. Consequently the potential along a line drawn through the central part of an opening in the grid structure has very little dip, as shown in *b*, Fig. 4-24. Electrons will, therefore, flow along this latter path even with the grid at a negative potential, maintaining the ionization of the gas in the tube and so preventing any reduction of plate current.

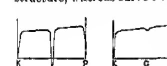


FIG. 4-24. Potential distribution inside a thyratron when passing plate current. Potential curves are taken along the dotted line shown in the lower part of each figure.

Obviously, if the grid was made so negative that the positive-ion sheaths on opposite sides of the opening overlapped, it would be possible for the grid to interrupt the flow of plate current. In the usual type of thyratron this potential would be so negative as

to draw a very excessive positive-ion grid current and is never applied. Thyratrons may, however, be constructed with a grid consisting of a very fine mesh, so that no path between cathode and anode can be found that does not pass very close to a portion of the grid surface. A grid of this type is capable of actually stopping the flow of current to the plate when made moderately negative but tends to draw a large current to itself. It will draw appreciable current when negative as well as when positive, owing to the continual presence of a small amount of ionization from cathode emission, so that considerable power is required to actuate this type of tube. It is, therefore, not widely used.

Type of Grid Structure Required. The construction of the grid in a thyratron must be such as to prevent any electronic flow to the plate whatsoever, no matter how small, when the tube is in the nonconducting state. If even a few electrons can get past the grid, so few as to constitute an inappreciable current flow, ionization may set in and the positive ions will immediately nullify the effect of the grid, permitting full current to flow. Consequently, it is essential to the successful operation of the tube that the grid so completely shield the cathode from the anode that, with the grid more negative than the starting voltage, no electrons can reach the anode.

In high-vacuum triodes, on the other hand, the amount of current flowing at any one time is dependent solely on the number of electrons that are able to overcome the effect of the negative charge on the grid, either by a high initial emission velocity or by traveling around the grid. Normally there is a small current flowing in these tubes even when the grid is highly negative, as some electrons are emitted in such a direction and at sufficiently high velocities to travel around the grid and so come within the field of the plate. This current is too small to be observed with any but the most sensitive instruments and so does not materially affect the operation of the tube; but if it should exist in a thyratron, ionization would set in and the tube would pass full current at all times without regard to the voltage on the grid. The grid of a thyratron is therefore so made as to enclose the cathode completely. This accounts for its peculiar construction and large size, as seen in Figs. 4-22 and 4-32.

Another reason for a grid of such large proportions is that it must not emit electrons, as a flow of electrons from grid to anode

would fire the tube no matter what might be the grid potential. A certain amount of active material gradually deposits on the grid from the cathode during operation of the tube, so that the grid must have sufficient radiating surface to keep its temperature below that which will produce emission from the cathode material.

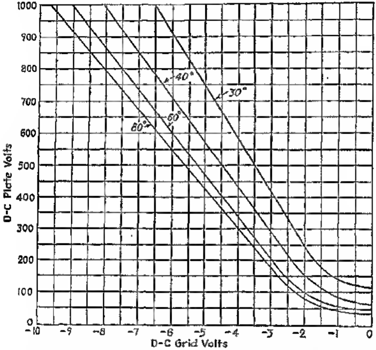


FIG. 4-25. Control characteristics of a small, negative-grid thyratron.

Thyratron Characteristic Curves. The characteristic curves of a small thyratron are shown in Fig. 4-25. Any combination of grid and anode potentials and temperature that determines a point to the right of a characteristic curve indicates that the tube will conduct; but if the point determined is to the left of the curve, the tube will not ignite upon application of the anode voltage if the grid voltage has been applied first. The temperature indicated on the curves refers to that of the condensed mercury; thus the tube

will break down at lower anode voltages as the temperature is increased.

The thyratron of Fig. 4·25 has a rated deionization time of about 1000 μsec For service with d-c anode supply (as in inverter service, page 239) a shorter deionization time is desirable. Such a tube may be constructed by locating the grid baffle very close to both the anode and the cathode, decreasing the size of the opening in the baffle, and decreasing the spacing between the glass envelope and the upper part of the grid cylinder and top of the anode, although the close spacing between grid and cathode results in a higher preare grid current.

The characteristic curves of Fig 4·26 are illustrative of the performance of a small tube of this type having a deionization time of about 100 μsec It may be seen that the curves lie more toward the positive region; thus this tube is known as a *positive-grid-control tube* in contrast to the *negative-grid-control tube* of Fig. 4·25

Use of Gases Other than Mercury Vapor. Mercury vapor is used in most tubes because the gas pressure is adjustable within the desired limits by temperature control of the liquid mercury in the bottom of the bulb. On the other hand this very characteristic is a source of trouble in certain applications where either (1) the temperature of the surrounding atmosphere varies between wide limits (as in outdoor installations) or (2) it is necessary that the tube characteristic remain constant within rather close limits (as in cathode-ray sweep circuits, page 306). Inspection of Figs. 4·25 and 4·26 shows that only a small variation in temperature will cause mercury-vapor thyratrons to break down at an entirely different anode voltage for the same applied grid voltage.

The temperature problem may be solved by using an inert gas, such as helium, argon, or neon, in the tube, usually argon. A tube so filled has a characteristic that is nearly independent of temperature and thus will always break down at approximately the same anode voltage Another advantage of using an inert gas is that the positive ions are lighter than those of mercury, permitting the tube to deionize more rapidly. Thus tubes containing these gases a[illegible]etter suited to higher frequency operation than are mercur[illegible]or tubes. Even argon-filled tubes, however, are generally [illegible]tisfactory for service with currents of a frequency of much more than a few thousand cycles.

The disadvantages of using an inert gas are that the tube drop is about 50 per cent greater than with mercury vapor, and the inverse breakdown voltage is much lower. Thus argon-filled tubes

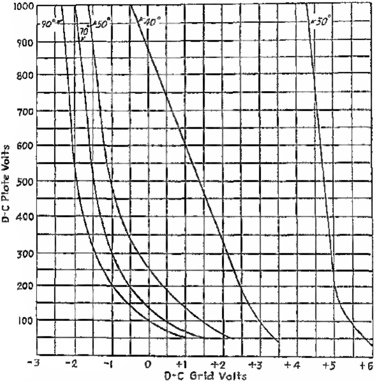


FIG. 4·26. Control characteristics of a small, positive-grid thyratron.

are normally used only in those applications where mercury-filled tubes are not satisfactory.

Comparison of the Thyratron with the High-vacuum Triode. It is obvious from the difference in the nature of the grid-control characteristics that the fields of use of the thyratron and the high-vacuum tube do not materially overlap. The thyratron is essentially an electronic switch, having an internal impedance that is either very high (switch open) or very low (switch closed) depend-

ing on the grid potential, although it cannot be opened by the same means as it was closed, *viz.*, the grid.[1] The high-vacuum tube, on the other hand, has more of the characteristics of a valve, being able to vary the magnitude of the current throughout a wide range. The current capacity of the thyratron is much higher than that of the high-vacuum tube, and the plate voltage necessary to produce this current is much lower. Thus the thyratron is essentially a power device, while the high-vacuum tube is widely used in communication circuits, for controlling thyratrons, and in other low- and medium-power applications.

Grid Control of Mercury-arc Tubes. Grids may be inserted in mercury-arc tubes to perform the same functions as in thyratrons. Since most mercury-arc tubes have more than one anode, they must also have more than one grid, each grid being located in the path of the electrons flowing to a particular anode. The performance of such a unit follows essentially the same laws as the thyratron, and the same precautions must be observed in the design of the grids to insure proper controlling action.

Grid Control of Ignitrons. Grid control is not applied to ignitrons directly, but it is possible to apply grid control to the tube that fires the igniter circuit. This tube may be a thyratron capable of determining the instant of time at which ignition of the arc in the ignitron is to take place. (For circuits see Chap. 8.)

Grid-controlled Cold-cathode Tubes. Cold-cathode tubes are also constructed with a third, or controlling, electrode. This third electrode is often referred to as a grid by analogy with the high-vacuum tube, although it usually consists of a small wire. One such tube is shown in Fig. 4-27 and in cross section in Fig. 4-28. By adjusting the grid to a more negative potential than a certain critical value—different for each value of plate potential—the few ions originally present may be prevented from reaching the anode, and no current will flow through the tube. This is illustrated by the curves of Fig. 4-29 where the area to the left of the curves represents combinations of plate and grid voltages that will not cause the tube to become conducting, whereas the area to the right represents the opposite effect.

If the grid of one of these tubes is left free, a plate potential of several hundred volts is necessary to break it down; but if the plate

[1] Actually the magnitude of an *alternating current* may be controlled by special circuits (p. 232), but with direct currents no such control is possible.

and grid potentials are adjusted to a point just a little to the left of the appropriate curve of Fig. 4-29, the tube will be very sensi-

Fig. 4-27. Cold-cathode, grid-glow tube. (*Westinghouse Electric Corp.*)

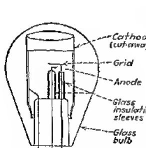


Fig. 4-28. Sketch of the principal parts of the tube of Fig. 4-27.

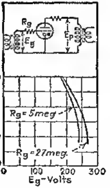


Fig. 4-29. Typical curves of a tube such as that of Fig. 4-27.

tive to slight changes in grid voltage; a man's hand placed near a wire connected to the grid will, for example, cause the tube to break down. Thus the tube may be used to close a relay to indicate the presence of an undesired person, such as a burglar, or to detect the presence of metallic objects on a conveyor belt or in a chute, etc. A typical circuit is given on page 250.

Another tube of this same type is shown in Fig. 4-30.[1] The large disk is the cathode, and the anode consists of a small wire protruding through the center of the cathode. The grid, or starter, as it is called in this tube, consists of a small circle of wire located quite close to the cathode. The tube depends for its operation on the application of Paschen's law, which states that the breakdown potential decreases—between certain limits—as the electrodes are brought into closer proximity. Thus the breakdown potential between the starter and the cathode is less than that between anode and cathode. A small potential applied between starter and cathode will thus produce current flow between these electrodes, and the resulting increase in ionization will cause the main anode to pass current at a lower potential than if such ionization were not present. A potential is normally applied between anode and cathode that is less than the breakdown potential between these two electrodes but that exceeds the breakdown potential when the starter is energized. Thus no current will flow through the anode circuit until a suitable voltage is applied to the starter. After the discharge forms between anode and cathode, the current will, of course, continue to flow, regardless of the starter potential, until the plate circuit is opened or its potential is lowered sufficiently. If a-c power is supplied to the plate, the plate potential is automatically removed at the end of each positive half cycle, and the

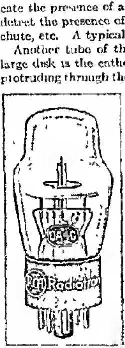


FIG. 4-30. Grid-controlled cold-cathode tube (*RCA*.)

[1] W. E. Bahls and C. H. Thomas, A New Cold-cathode Gas Triode, *Electronics*, **11**, p. 14, May, 1938.

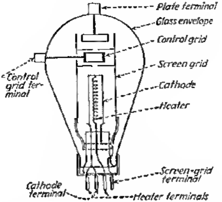


FIG. 4-31. Cross-sectional view of a small, screen-grid thyratron.

grid will then resume control. (This is similar to the operation of thyratrons with a-c supply, see Chap. 8.)

This tube is used primarily in control circuits where stand-by service is desired. No power is consumed by the tube until the starter electrode is energized, but sufficient current will flow from anode to cathode after firing to operate a reasonably rugged relay. A circuit and applications are given on page 249.

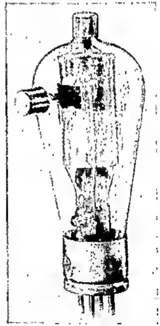


FIG. 4-32. A small, screen-grid thyratron. (*General Electric Co.*)

3. SCREEN-GRID TUBES

All thyratrons of the single-grid type tend to draw an appreciable grid current, even when the plate is not conducting. Grid-current flow during periods of plate conduction is to be expected, but even during nonconducting periods a small amount of ionization is caused by the emission of electrons from the cathode, and

these ions are attracted to the negative grid. This grid current may be greatly reduced by the insertion of a second, or screen, grid located between the control grid and the cathode and anode.

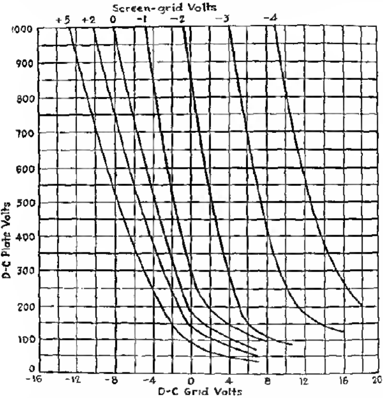


FIG 4-33 Control characteristics of a small, screen-grid thyratron

The internal construction of a small screen-grid thyratron is shown in Fig. 4-31, and a photograph of a complete tube in Fig 4-32. The screen grid is similar in construction to the grid used in triode-type thyratrons except that it has two baffle plates. The control grid is a small cylinder located between these two baffles. Thus the control grid is so located that cathode material cannot

readily deposit on its surface nor will it absorb appreciable heat from the cathode or anode or from the arc stream. Its small size greatly reduces the positive-ion current flowing to it when it is negative and also reduces its electrostatic capacity to the other elements. The grid current is, therefore, very small, and the tube may be used with a very high resistance control source such as a photoelectric tube.

The characteristics of this tube are given in Fig. 4-33 where it will be seen that the starting voltage may be made either positive or negative by adjusting the voltage on the screen grid.

Problems

4-1. A certain mercury-vapor diode requires 5 amp at 5 volts to heat the cathode. The tube drop at 1 amp plate current is 12 volts. (*a*) What is the total power lost in the tube at a current of 1 amp? (*b*) Approximately what is the total loss in the tube at 2 amp plate current? (Emission current in excess of 10 amp.)

4-2. A high-vacuum tube draws a filament current of 52 amp at 22 volts with a tube drop of 800 volts at 1 amp plate current. (*a*) What is the total loss in the tube at 1 amp plate current? (*b*) Assuming that the plate-current-plate-voltage characteristic is the same as for infinite parallel-plane electrodes, what is the loss at 2 amp? (Compare results with those of Prob. 4-1.)

4-3. A tungar tube requires a filament current of 14 amp at 2 volts with a tube drop of 10 volts at a load current of 5 amp. What is the total loss in the tube?

CHAPTER 5

PHOTOELECTRIC TUBES[1]

Photoelectric tubes fall into three general classes: (1) photoemissive tubes, (2) photoconductive cells, and (3) photovoltaic cells. Photoemissive tubes are those in which impinging light causes emission of electrons from a photosensitive surface. Most photoelectric devices fall in this class. Photoconductive cells are those in which the resistance varies with the intensity of the light. Photovoltaic cells are those which generate an internal emf upon being exposed to light.

Units Used in Light Measurement. Before going further it may be well to consider briefly some of the units used in measuring light intensities. Unfortunately these are none too familiar to many engineers.

Candle Power. The *international candle* is the unit of light intensity. Electric lights are rated in candle power according to the intensity of the light that they emit.

Lumen. The lumen is a unit of the luminous flux emitted from a light source. If a uniform point source of light, having an intensity of 1 cp, is located at the center of a sphere of 1-ft radius, the amount of luminous flux on 1 sq ft of the sphere's surface is called a *lumen*. Evidently a 1-cp source emits a total of 4π lumens.

Foot-candle. The foot-candle is a measure of illumination, or luminous flux density. It is equal to a lumen per square foot.

1. PHOTOEMISSIVE TUBES

Commercial photoemissive tubes (commonly abbreviated to *phototubes*) consist of a cathode and an anode, contained in a glass tube very similar to those used for thermionic vacuum tubes. The cathode generally consists of a half cylinder of copper or silver coated with a photosensitive material, with a vertical wire anode lying along the center line of the cylinder (Fig. 5-1). The cathode

[1] It is suggested that the reader review the sections on photoelectric emission in Chap. 1.

is purposely made with a comparatively large surface, to secure the maximum possible emission. On the other hand, the small magnitude of the currents obtained even with fairly large cathodes makes a large anode entirely unnecessary. Evidently a large anode would greatly reduce the amount of light striking the cathode and so lessen the sensitivity of the tube.

Phototubes, like thermionic tubes, are of two types: high-vacuum and gas-filled. The gas-filled tubes ionize when the plate voltage exceeds a certain value and thus pass a much larger current than do the high-vacuum tubes. Unlike the thermionic tube, however, this ionization is not produced for the purpose of neutralizing space charge, since, as will be shown, there is no limitation of current flow by space charge under normal operation. The ionization of the gas increases the current flow by liberating positive ions and negative electrons from the ionized gas atoms which then flow to cathode and anode, respectively, the positive ions drawing additional electrons from the cathode.

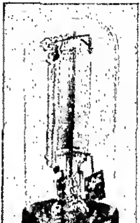


FIG. 5·1. Westinghouse SR-50 gas-filled phototube.

High-vacuum tubes are used in light-measurement work and in certain relay-operating applications. They are less subject to damage due to applications of excess voltage or current, and their sensitivity remains more constant over a period of time. Gas-filled tubes are used quite largely in talking-picture work where their higher sensitivity reduces the amplification needed.

Characteristic Curves. High-vacuum Tubes. The curves of plate current vs. plate voltage for a high-vacuum phototube are very similar to those for a high-vacuum diode. The magnitude of the current flowing under any given condition is determined by the electric gradient of the space charge surrounding the cathode and the velocity of emission of the electrons just as in the hot-

cathode diode. One very marked difference, however, is that the cathode of the diode is small in area, causing relatively high negative gradients, whereas that of the photocell is large in area producing much lower negative gradients at the surface. Thus the phototube plate current saturates at lower plate voltages, as the electrons encounter a lesser retarding force from the space charge Figure 5-2 shows a family of such curves. The current is seen to rise rather abruptly with increasing potential until a point of saturation is reached. This saturation point is determined primarily by the emission and therefore varies with the amount of light applied to the tube. Evidently the phototube is normally operated on the saturated part of its characteristic curve, since it is desired that the current be under the control of the incident light rather than of the applied plate potential. This is, of course, just the reverse of the thermionic diode, where the emission is held constant and the current flowing is controlled by the potential applied between plate and cathode

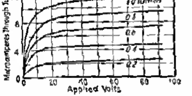


Fig 5-2 Current-voltage characteristics of a high-vacuum phototube

If points are taken from the curves of Fig 5-2 for any potential well above saturation and plotted against light intensity, the result will be the curve of Fig. 5-3. This curve is seen to be a straight line passing through the origin, so that the current flowing through the cell is a direct measurement of the light striking its surface. It should also be evident that this curve is virtually independent of the applied potential, unless this potential is allowed to fall below the saturation point Thus, for all practical purposes, there is but a single curve in this family, just as there was but a single emission curve for the hot-cathode vacuum tube.[1]

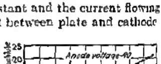


Fig 5-3. Current-light characteristic of the tube of Fig 5-2 for an anode voltage of 90

[1] Very low plate voltages (below about 40 volts for this tube) will produce a curve somewhat below the one shown, since saturation current will not flow at low plate voltages

Characteristic Curves. Gas-filled Tubes. Gas inserted into a phototube will increase its sensitivity and otherwise alter its characteristics. Figure 5-4 illustrates this difference quite vividly. The curves for both the vacuum and the gas-filled tubes virtually coincide up to a plate potential of about 20 volts. At that point ionization of the gas sets in, and the current increases fairly rapidly from that point on. Evidently the maximum current of the gas-filled tube is greatly in excess of the total photoemission of the cathode.

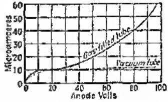


FIG. 5-4. Curve showing the effect of gas on the 1.0 lumen characteristic curve of Fig. 5-2.

The curve of Fig. 5-4 is essentially that of Fig. 4-2 from *o* to *b* (except for the interchange of abscissa and ordinate). The region *o* to *a* of Fig. 4-2 corresponds to the operation of the high-vacuum phototube, although the source of emission was ionization by cosmic rays instead of photoemission, while the region *a* to *b* corresponds to the region above about 20 volts in Fig. 5-4 for the gas-filled tube. If the voltage impressed on a gas-filled tube is increased sufficiently, the breakdown potential of the gas will be

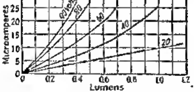


FIG. 5-5. Current-light curves of a gas-filled phototube. Note that the curve for low voltage (where there is no ionization) is similar to that of Fig. 5-3.

exceeded and a glow discharge may be established, corresponding to the region *c* to *f* in Fig. 4-2. The resulting increased positive-ion bombardment is extremely injurious to the cathode, and care should be exercised to avoid application of sufficient potential to cause breakdown.

Figure 5-5 shows a family of curves for the gas-filled tube, with impinging light as the abscissa. It will be noticed that the current

is no longer independent of the plate potential but increases continuously as the potential is raised. It should be noticed further that at low plate voltages the curve is straight just as for the high-vacuum tube, but at higher voltages the plate current increases at a power higher than the first. If this type of tube is used to reproduce variations in light, as in talking-picture applications, a certain amount of distortion will result, *i.e.*, the current flow in the tube will not be directly proportional to the impressed light. However, inserting a sufficiently high load resistance in series will reduce the distortion to a negligible quantity,[1] and therefore, gas-filled tubes are widely used in talking-picture work, being preferred to the high-vacuum type because of their greater sensitivity.

Another limitation to the use of gas-filled tubes is that they do not respond so rapidly to variations of light intensity as do the high-vacuum tubes. Thus they tend to give a lower response when the light varies at, say, 5000 cycles/sec than when it varies at only 100 cycles/sec. In talking-picture applications this will also cause distortion of the reproduced sound unless it is corrected by suitable electrical networks. Modern gas-filled tubes, however, have a response that is sufficiently fast to cause only a very small amount of distortion in the frequency range required.

The gas used in photoelectric cells is always one of the inert group, since the photosensitive surfaces are all rather highly active chemically. Argon is most commonly used, being relatively cheap and entirely satisfactory in other ways.

Color Response of Phototubes. Most photoemissive materials respond more readily to the violet and ultraviolet end of the spectrum than to the red and infrared, as illustrated by the curves of Fig. 1-7 (page 21), although, as stated on page 21, certain composite materials are capable of a much wider spectral response. Figure 5-6 is illustrative while Fig. 5-7 shows the response of another cesium oxide tube wherein the response is peaked in both the

[1] See discussion of distortion in thermionic tubes in Chap. 9. Briefly the curvature of the characteristic curve may be considered as being due to a variation in the internal resistance of the tube. If a constant resistance is connected externally in series with the tube, which is much larger than the equivalent tube resistance, it is evident that the magnitude of the current flow will be determined largely by the constant external resistance rather than by the variable internal resistance. The current will therefore be nearly proportional to the intensity of the impressed light.

red and ultraviolet regions.[1] The insertion of gas produces very little change in the color sensitivity.

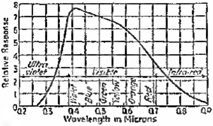


FIG. 5-6. Color response of a high-vacuum, cesium oxide phototube. This tube is largely responsive to the violet and ultraviolet end of the spectrum.

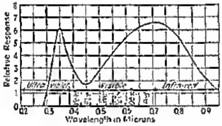


FIG. 5-7. Color response of a high-vacuum, cesium oxide phototube. The response of this tube has been peaked in the red end of the spectrum, as well as in the ultraviolet.

Phototube Circuits. The simplest type of circuit for use with a phototube is that of a battery and microammeter in series with the tube (Fig. 5-8). If a high-vacuum tube is used, the current read on the meter will be a direct indication of the light striking the tube and may be used as a measure of light intensity. The phototube is seldom used in this manner, however, since the current that it will pass is so small as to require a very sensitive meter; furthermore the photovoltaic cell, to be discussed in a later section, provides a much simpler means of accomplishing the same result.

The more common applications of phototubes involve the use of associated vacuum-tube amplifiers.[2] The tube is coupled to the

[1] See discussion of the cesiated-silver surfaces on p. 20.

[2] See pp. 251 and 310 for further details.

input tube of an amplifier by means of a high resistance R_g (Fig. 5-9) between grid and cathode. Since the current flow through the phototube is of the order of a few microamperes, this resistance should be very high, preferably several tens of megohms. For the same reason the insulation resistance within the phototube between its cathode and anode, and within the triode between grid and cathode, should be at least as high and preferably higher.

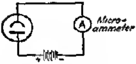


Fig. 5-8. Simple circuit for using a phototube to measure light intensities.

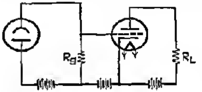


Fig. 5-9. Circuit of a photoelectric tube coupled to the input of a vacuum-tube amplifier.

Special provisions are often included in the construction of phototubes to ensure the resistance of the leakage path remaining high throughout the life of the tube.

2. PHOTOCONDUCTIVE CELLS

Many materials show the interesting property of changing electrical resistance when illuminated. The most commonly used of these is selenium, but lead carbonate, molybdenite, rock salt, cinnabar, and a few other substances show the same effect to a greater or lesser degree. The effect seems to be a true photoelectric phenomenon, but no electrons are actually emitted from the surface, the liberated electrons apparently flowing entirely within the material and recombining with the positive ions at a rate that eventually causes saturation.

Selenium Cells. Selenium cells are generally built on a grid of very fine wire (Fig. 5-10). Two wires are wound side by side on a suitable insulating support, such as porcelain or glass, the spacing between the wires being so close that it can hardly be discerned by the naked eye. Selenium is then distilled or painted onto this surface, following which it is annealed at a temperature close to its

melting point, to produce the gray, crystalline variety of selenium. It is only this variety of the metal which exhibits the photoconduc-

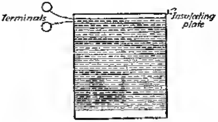


FIG. 5-10. Illustrating one method of building a selenium cell.

tive action; hence the annealing. The finished cell is then mounted in a glass container filled with an inert gas (Fig. 5-11).

The purpose of the grid construction in these cells is to secure an internal resistance that is not too high. Selenium is not a good conductor, and even with the type of construction just described the *light* resistance is of the order of 1 megohm when illuminated by 100 ft-candles of light, whereas the *dark* resistance[1] is from six to ten times this value. Evidently the amount of current that such a cell will pass with the normal impressed potential of about 100 volts is not much over 100 μa.

Another problem that the grid construction solves quite well is that of exposing a large amount of the selenium to the impressed light. Selenium is quite opaque, so that light cannot penetrate a great distance into the material. Excessive thickness of active material is therefore highly undesirable, since it will tend to decrease the dark resistance without decreasing the light resistance a proportionate amount. The ratio of dark to light resistance, the measure of a cell's efficiency, will therefore decrease with the thickness of the material.

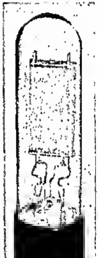


FIG. 5-11. General Electric FJ-31 selenium cell.

[1] By dark resistance is meant the resistance between terminals of the cell when it is not illuminated.

The current-light characteristic curve of the tube of Fig. 5-11 is shown in Fig. 5-12. It is not so linear as that of the high-vacuum photoemissive tube, but the current variation is much greater. If the light intensity is varied periodically, it will be found that the cell response drops considerably as the light modulation frequency is increased above a few hundred cycles per second.

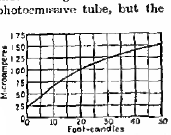


FIG. 5-12. Current-light response of a typical selenium cell.

The normal selenium cell responds readily to infrared and red radiation and gives only a feeble response to violet and ultraviolet, although cells may be manufactured with quite a variety of color-response characteristics by slight changes in manufacturing processes.

Other Types of Cells.[1] A thallous sulfide cell has been developed which has a maximum response at 9500 A and gives a satisfactory response as high as 14,500 A, well into the infrared region.[2] It has a resistance variation of from 0.1 to 20 megohms and a linear response for light intensities below 0.02 ft-candles, although a reasonably linear response may be obtained at very much higher intensities if the tube is operated in series with a high resistance. Its response falls off somewhat with increasing frequency of modulation of the light source but the signal-to-noise ratio is much higher than with cesium oxide phototubes. Thus, while not ideal for such services as reproduction of sound on film, it may be more satisfactory than a phototube when used with a suitably compensated amplifier.

At the other extreme of color sensitivity is the lead sulfide cell which has a maximum response at 2500 A and a threshold at 3600 A and is, therefore, primarily responsive to ultraviolet light. The response of this cell is nearly constant at light intensities up to 40 ft-candles. It has a resistance variation of from 0.1 to 20 megohms.

Circuits Suitable for Photoconductive Cells. Photoconductive cells pass sufficient current to operate a very sensitive relay directly. However, they are more commonly used with vacuum-tube ampli-

[1] See abstract in *Proc. Optical Soc. Am.*, 35, p. 356, June, 1946.
[2] See definition of angstroms (A) on p. 21.

fiers in circuits similar to those used with phototubes, as Fig. 5-9. For maximum response the resistance R_g should be the geometric mean of the dark and light resistances of the tube. In this they differ from the phototube which should be used with as high a resistance as possible.[1] The obvious reason for this is that the response of the latter type is due to a variation in its *current* (emission), whereas in the photoconductive cell it is due to a variation in *resistance*.

3. PHOTOVOLTAIC CELLS[2]

The photovoltaic effect was discovered by Becquerel about the middle of the nineteenth century. He found that, if two similar electrodes are immersed in an electrolyte and the electrolyte or one of the electrodes is illuminated, a small potential is set up between the two electrodes, which is a function of the intensity of the impinging light. This phenomenon is apparently due to emission of electrons from the illuminated electrode into the electrolyte. Light-sensitive devices operating on this principle are known as *photovoltaic cells*.

Photovoltaic cells are connected directly to a meter or other load without any external source of emf; in fact the use of an external source of emf will cause serious damage if not destruction of the cell. They are capable of producing sufficient current to operate a sensitive relay directly and are therefore extremely useful in control applications where a source of potential, such as is required for photoconductive cells or phototubes, is not available. The short-circuit current flowing is almost directly proportional to the intensity of the impinging light so that light-intensity measurements may be made by connecting a sensitive, low-resistance meter across the terminals of a cell, as in photographic, light-intensity meters.

Photovoltaic cells tend to be somewhat slow in responding to sudden changes in light. Thus if the intensity of the light is varied at a rate of several thousand cycles per second, the changes in

[1] Except that it should not be so high as to drop the voltage across the phototube below that which will ensure the flow of saturation current, nor is there any point in increasing the external resistance until it approaches the shunting resistance of the balance of the circuit.

[2] For some of the applications of this type of cell see Anthony Lamb, Applications of a Photoelectric Cell, *Elec. Eng.*, 54, p. 1186, November, 1935.

current through the cell will be much less in magnitude than if the light is varied at only a few cycles per second. The curve of Fig 5-13 illustrates this point, where the abscissa is the frequency at which the light is varied in intensity and the ordinate gives the effective current flow in percentage of maximum. Thus these cells are not suitable for use in communication circuits.

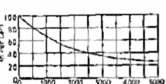


FIG 5-13 Dynamic-response curve of a photovoltaic cell

Iron Selenide Cell. The iron selenide cell (Photronic) consists of an iron electrode coated with iron selenide, this in turn being coated with a thin, translucent, metallic film for the other electrode. The complete unit is mounted in a suitable case with a glass window to admit light. Current-response curves for this type of cell are shown in Fig. 5-14. The color response is quite similar to that of the human eye.

Equivalent Circuit of Photovoltaic Cells. The curves of Fig. 5-14 show that the current produced by a short-circuited photovoltaic cell (3 ohms is virtually a short circuit) varies in a nearly direct proportion to the intensity of the impressed light whereas the current flowing with higher resistances is not at all a linear function of the light. This is because the cell is more nearly a constant-current generator than constant-voltage. The equivalent circuit of this type of cell is shown in Fig 5-15 in which G represents a generator, producing a current almost directly proportional to the intensity of the light, which, to a close approximation, maintains this current flow regardless of the external circuit conditions. The resistance R is internal to the cell and tends to decrease as the illumination is increased. The current flowing through the external circuit depends upon the ratio of external to internal resistance as well as upon the intensity of the light. As the external resistance is increased, the percentage of the current produced by G

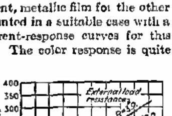


FIG. 5-14 Current-light curves of a Weston Photronic cell

that passes through R increases until under open-circuit conditions all the current flows through this shunt path. This resistance, for the Photronic cell, varies between about 7000 and 1000 ohms, depending upon the intensity of the light. Under short circuit, all the current generated of course flows through the external circuit, and a nearly straight-line relationship is obtained (Fig. 5-14).

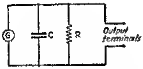


FIG. 5-15. Equivalent circuit of a photovoltaic cell. The generator shown delivers a current which is a function of the impressed light alone.

FIG. 5-16. Circuit using a photovoltaic cell to control a load which may consist of a relay, meter, or other device.

Circuits Suitable for Use with Photovoltaic Cells. Photovoltaic cells are most commonly used directly in series with a relay, meter, or other load, as in Fig. 5-16. Since they generate their own internal emf, no external source is required, and the circuit is instantly ready for use without power consumption from an external source. Photoelectric tubes, on the other hand, must normally be used in connection with vacuum-tube amplifiers which must consume power continuously for cathode heating and plate-current supply to be ready for service on demand. Photoconductive cells also require consumption of power at all times. Thus the photovoltaic cell supplies a very real demand for control of various electrical and mechanical circuits and devices with a maximum efficiency and a minimum of auxiliary apparatus.

Photovoltaic cells cannot be used satisfactorily with vacuum-tube amplifiers, since they produce only a very low voltage, being essentially constant-current generators. For operation of a vacuum-tube amplifier, a high resistance must be inserted in series with the cell to build up the voltage required by the grid of the tube, and, as shown by Fig. 5-14, insertion of resistance will materially reduce the current flowing. It may also be seen that the maximum voltage obtainable with 10,000 ohms resistance is about 0.25 volt. It should be further evident that an increase in resistance above

this amount will result in very little further increase in voltage. Since a good photoelectric tube is capable of producing an output of at least several volts with a reasonable intensity of light, it is evident that the photovoltaic cell is not suitable for such service.

Problems

5-1. A phototube having the characteristics of Fig. 5-3 is to be used in the circuit of Fig. 5-9. The constants of the amplifier tube are $\mu = 13$, $r_p = 13{,}000$ ohms, $g_m = 1000$ μmhos. If a light flux of 0.2 lumen is impressed on the phototube, what will be the change in plate current in the amplifier tube as the light is cut off when R_s is (*a*) 0.1 megohm, (*b*) 1 megohm, (*c*) 5 megohms? (Let $R_L = 0$ and assume μ, r_p, and g_m to remain constant.)

5-2. A photovoltaic cell having the characteristics of Fig. 5-14 has an active exposed area of 2 sq in. What current will flow through a 3-ohm load with 1.0 lumen of light flux on the cell?

CHAPTER 6

SPECIAL TYPES OF TUBES

A number of tube types do not fit into the classifications thus far considered. Some of these are cathode-ray tubes, X-ray tubes, iconoscopes, kinescopes, and electron multipliers. The most important of these special tubes are briefly described in this chapter. No attempt is made either to make the list complete or to give a thorough discussion of those which are described, since the number and variety of tubes which are available today are too extensive to make such a procedure either possible or desirable. Nor is any attempt made to describe the many tubes developed for ultra-high-frequency operation, such as the magnetron and klystron, since this field has become too extensive to be adequately covered in a single book along with the fundamentals of the performance of the more conventional types of tubes suitable for use at the lower frequencies.

Cathode-ray Tubes. The cathode-ray tube normally consists of six parts contained within an evacuated bulb: (1) a hot cathode for producing the electrons, (2) a grid to control the intensity of the beam, (3) a first anode to draw the electrons from the cathode into the beam, (4) a second anode to accelerate the electrons and to focus them on to the screen, (5) electrodes or coils to deflect the electron beam by means of the emf or current to be observed, and (6) a fluorescent screen. The first four parts constitute what is often termed an *electron gun*, since they serve to generate and direct the stream of electrons that constitutes the moving element of the oscillograph. The electron gun may also be used in other applications such as the orthicon and kinescope (page 154).

Figure 6-1 shows the principal parts of a modern cathode-ray tube.[1] A is the cathode, usually of the indirectly heated type, such as was described in Chap. 1. Only the end is coated, since

[1] See also Fig. 6-11 (p. 156) showing a cutaway view of a gun used in an iconoscope.

emission from that part of thimble A alone is useful in forming the electron beam.

B is the grid for controlling the intensity of the beam. It is not made in the form of a mesh or screen as in the triode tube but is a solid metal sleeve having a cylindrical disk immediately in front of the cathode. A small aperture in the center of this disk permits the passage of those electrons which may be traveling in the right direction. The potential of the grid is made negative with respect to the cathode, and the amount of this negative potential determines the number of electrons in the beam and therefore its intensity.

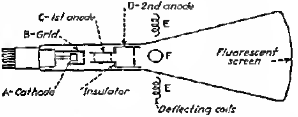


FIG. 6-1. Sketch showing the principal parts of a cathode-ray tube.

C is the first anode and is made moderately positive with respect to the cathode, the exact potential depending upon the size of the tube. It is very similar to the grid in construction, having one or more small perforated disks in the path of the electron stream. The positive potential of this anode draws electrons through the grid aperture into the beam, while the small openings in its disks, known as the *masking apertures*, cut off some of the peripheral electrons, much as the stop in a camera cuts off some of the light. In fact, the gun as a whole may be analyzed as an electron optical system by considering the electric fields as thick lenses.[1]

Beyond the second aperture the beam enters the field of the second anode D, which is maintained at a potential of from a few hundred volts in small tubes to several thousand volts in large ones. The field set up by this anode, besides providing further acceleration, gives a radial velocity to the electrons of the beam,

[1] Maloff and Epstein, Theory of the Electron Gun, *Proc. IRE*, 22, p. 1386, December, 1934.

acting toward the axis, such that, if the axial velocity is correct, all electrons will concentrate in a single spot on the fluorescent screen at the end of the tube. The axial velocity is, of course, also affected by the potential of the first anode, so that the spot may be adjusted to any size at will by varying the potential on anode C. This makes it unnecessary to vary the high-potential source for focusing and greatly facilitates control of the spot on the screen.

Deflecting coils are shown at E and F, coils E deflecting the beam in a horizontal direction while coils F act in a vertical plane. Passage of the current to be investigated through one pair of coils, and a sweep or timing current through the other, will produce an image of the unknown current on the screen. Circuits for producing the sweep action are given in Chap. 9 (page 304).

Deflection of the electron stream may be accomplished by electrostatic as well as magnetic means (Fig. 6-2). Two pairs of plates are inserted in the tube at right angles to each other and so situated that the electron beam must pass through the field set up between each pair after leaving the second anode. The beam may then be deflected by the application of emfs, the only current drawn being that due to the capacitance between the deflecting plates.

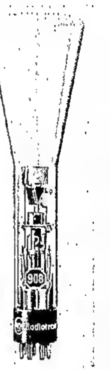


FIG. 6-2. Photograph of a cathode-ray tube showing details of gun structure and location of beam-deflecting plates. (*RCA*.)

Electrostatic deflection is usually preferred in small tubes operating at relatively low anode potentials. Large tubes require much higher anode voltages which materially reduce the deflection sensitivity. For these tubes magnetic deflection is often preferred since the sensitivity is then less affected by an increase in anode voltage. On the other hand magnetic deflection is open to the objection that positive ions, which move more slowly than the electrons, are not appreciably affected by the deflecting field and will always fall on the center portion of the screen, tending to damage the

fluorescent material and produce a dark spot. Another objection to magnetic deflection is that, at the higher frequencies, it is difficult to secure sufficient field intensity to produce satisfactory deflection.

The sensitivity of a tube with electrostatic deflection may be improved by inserting a conducting ring, known as an *intensifier electrode*, around the inside wall of the tube at a point close to the screen, and by applying a high potential between this ring and the cathode. This permits the use of lower potentials on the first and second anodes with resulting improvement in sensitivity yet provides the high electron velocity required for good intensity by

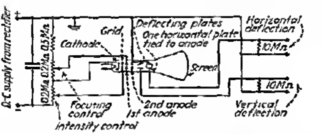


FIG. 6-3. Circuit of a cathode-ray tube showing the method of controlling the focus and intensity.

accelerating the electrons after they have been deflected. The improvement in sensitivity under this arrangement is due to the lower velocity of the electrons at the time they pass between the deflection plates.

Electrons, after striking the screen, must be returned to the cathode, else a negative charge would build up on the screen which would stop the beam. In modern tubes this return current is provided by secondary emission from the screen which is sufficiently conducting for this purpose. These secondary electrons are attracted to the nearest anode whence they are returned to the cathode through the external circuit.

The circuit diagram of an electrostatically controlled tube and its associated circuit is shown in Fig. 6-3. The positive side of the supply is grounded rather than the negative, to keep the screen end of the tube at ground potential. The electrodes of the gun are indicated symbolically only, their actual construction being as in Fig. 6-1. One deflection plate of each pair is tied to the common

ground terminal so that care must be exercised to avoid short circuits in making external connections. In some tubes all four plates are brought out to separate terminals, but care must then be used to avoid errors due to stray fields when neither plate of a pair is grounded.

Oscilloscopes. Cathode-ray tubes are often combined in a single unit with two high-gain amplifiers (one for each pair of plates), a suitable sweep circuit and a rectifier to supply the required direct potentials. Such combinations, known as *oscilloscopes*, make possible the reproduction of the wave shape of currents and voltages on the screen of the tube where they may be viewed or photographed. Oscilloscopes are built to reproduce currents having frequencies as high as several megacycles per second, which means that they are also capable of reproducing rather high-speed transients. A typical circuit is given on page 304.

As previously stated, the intensity of the spot on the screen is varied by changing the potential of the control grid while focusing is controlled by adjusting the potential of the first anode. Unfortunately these two controls are not entirely independent of each other and, for best results, any change in one must ordinarily be accompanied by a small change in the other. When the tube is used in an oscilloscope for observing the wave shapes of electric currents and potentials, this adjustment may be easily made if desired, but for television and other similar uses it is important that the focus be independent of any changes in beam intensity. This problem may be solved by inserting a third anode and maintaining the first and third anodes at a fixed potential, securing proper focusing by varying the potential of the second anode. The potential of the third anode is higher than that of the other electrodes so that the potential of the second anode may be made low enough to permit satisfactory control of the beam intensity by a potentiometer. Such an arrangement permits the first anode to function in much the same manner as the screen grid in a pentode, preventing changes in the field set up by the second anode from affecting the field around the grid.

Different types of screen material are available, some having a short persistence, such as might be needed in television where one picture must be replaced with another every thirtieth of a second, while others have a relatively long persistence, thereby aiding in the reproduction of transient phenomena. With a long-persistence

screen the trace of the transient will remain on the screen for an appreciable period of time permitting visual inspection or photographing. For most laboratory use zinc orthosilicate is used, giving a yellow-green color to which the eye is particularly sensitive and having a moderately long persistence. Television tubes have a white screen with moderately short persistence while very short persistence screens are commonly bluish in appearance.

Electron-ray Tube. An interesting variation of the cathode-ray principle is found in the electron-ray tube shown in a cutaway view in Fig. 6-4. The tube consists of two parts: a triode unit

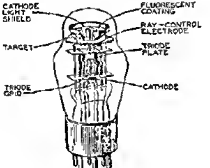


Fig. 6-4. Electron-ray tube.

and a fluorescent target mounted vertically above the triode portion. Electrons from that portion of the cathode within the bowl of the target are attracted to the bowl and cause fluorescence just as in the cathode-ray tube. The extent of the fluorescence, however, may be controlled by placing a third electrode between the cathode and target. In Fig. 6-4 this ray-control electrode is seen as a projecting pin just to the right of the upper cathode and welded to a projection of the plate of the triode unit.

The electron-ray tube is used as an indicator, *e.g.*, to indicate balance in a Wheatstone bridge or resonance in a radio receiver. A suitable circuit is shown in Fig. 6-5 where the control voltage is any suitable d-c source, such as the automatic-volume-control circuit of a radio receiver (see page 387). The grid of the triode unit is connected to the control voltage, so that as this voltage becomes more negative the plate current of the triode unit will

decrease. This will decrease the negative voltage between the ray-control electrode and the target, which will, in turn, decrease the shaded area. The amount of shaded area is, therefore, an indication of the magnitude of the d-c voltage. The shaded areas of Fig. 6-6 show the relative illumination for two different ray-control-electrode potentials, the potential being measured between this electrode and the target.

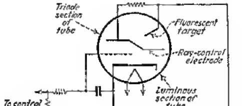


FIG. 6-5. Circuit for the electron-ray tube.

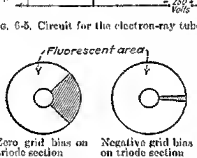


FIG. 6-6. Illustrating the change in fluorescent area of the electron-ray tube as the potential of the ray-controlling electrode is varied relative to that of the fluorescent screen.

Electron Multiplier. The electron multiplier is a device in which the phenomenon of secondary emission is utilized to produce amplification of small alternating emfs. Figure 6-7 shows the principle of operation. Electrons, emitted from a cathode (usually photoemissive), are attracted to a higher potential electrode P_1. Each of these electrons upon striking the surface knocks off several other electrons which are then drawn to an electrode of still higher potential P_2. This process may be repeated several times, thus resulting in a greatly increased current flow to the last electrode P_6. If each impinging electron knocks off N secondary electrons

and there are n multiplications or stages, the final current will be N^n times as large as the original cathode emission.

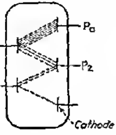


FIG. 6-7 Illustrating the principle of operation of the electron multiplier.

The electron multiplier of Fig. 6-7 is actually inoperable, since electrons from the cathode tend to flow directly to P_0, the plate of highest potential, and virtually no multiplication results. It is, therefore, necessary to provide a means of focusing the emitted electrons on the desired plate. One method of so doing is to provide a combination of electric and magnetic fields at right angles to each other, thus causing the electrons to travel in curved paths.

Figure 6-8 shows a tube utilizing this method of focusing.[1] Electrons emitted from the photocathode K are attracted toward the directing electrode D_1. A magnetic field is provided at right

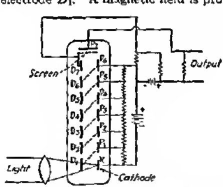


FIG. 6-8 General method of supplying the proper potentials to the several anodes of an electron multiplier with magnetic focusing. A magnetic field is set up normal to the plane of the page.

angles to the axis of the tube, causing the electrons to follow a curved path and thus strike the plate P_1. Secondary electrons from this plate are similarly directed to plate P_2 by the combined

[1] V. K. Zworykin, G. A. Morton, and L. Malter, The Secondary Emission Multiplier—A New Electronic Device, *Proc. IRE*, 24, p. 351, March, 1936.

action of the electric field between P_1 and D_2 and of the magnetic field. Since these tubes are used to amplify only very small initial emission currents, the current flowing to any electrode is small and the electrode voltages are easily supplied by a voltage divider as shown. The directing electrodes D_1, D_2 . . . are connected internally to the corresponding plates P_1, P_2 . . . as indicated by the dotted lines.

It is also possible to focus the electrons by purely electrostatic means, thus making unnecessary the application of an external field. The reader is referred to the literature for further details.[1]

Electron multipliers have proved useful in television where they have been used to amplify the low output of television pickup tubes and have even been incorporated into such tubes (see next section). They have also been used in the detection of light of very low intensity in such applications as astronomical photometry[2] and Raman spectrography.[3]

Special Tubes Used for Television.[4] For television purposes, a number of special tubes have been developed, among them the kinescope, the iconoscope, and the electron multiplier.

The kinescope (Fig. 6-9) is the tube used for reproducing the television image at the receiver.[5] Essentially it is a cathode-ray tube in which the incoming signal is impressed on the grid that

[1] V. K. Zworykin and R. A. Rajchman, The Electrostatic Electron Multiplier, *Proc IRE*, 27, p. 558, September, 1939; L. Malter, The Behaviour of Electrostatic Electron Multipliers as a Function of Frequency, *Proc. IRE*, 29, p. 587, November, 1941.

[2] G. E. Kron, Application of the Multiplier Phototube to Astronomical Photoelectric Photometry, *Astrophys. J.*, 103, p. 306, May, 1946.

[3] D. H. Rank and R. V. Wiegand, A Photoelectric Raman Spectrograph for Quantitative Analysis, *J. Optical Soc. Am.*, 36, p. 325, June, 1946.

[4] It is not within the scope of this book to present more than a few general statements concerning these very specialized tubes. For further details the reader is referred to books on television such as Donald G. Fink, "Principles of Television Engineering," McGraw-Hill Book Company, Inc., New York, 1940.

[5] For further information see V. K. Zworykin, Description of an Experimental Television System and the Kinescope, *Proc. IRE*, 21, p. 1655, December, 1933; Iconoscopes and Kinescopes in Television, *RCA Rev.*, 1, p. 60, July, 1936; R. R. Law, High Current Electron Gun for Projection Kinescopes, *Proc. IRE*, 25, p. 954, August, 1937; I. G. Maloff, Direct-viewing Type Cathode-ray Tube for Large Television Images, *RCA Rev.*, 2, p. 289, January, 1938.

controls the intensity of the electron stream and thus varies the intensity of the light on the viewing screen. The point of the screen on which the electron stream is focused at any given instant of time is controlled either electrostatically as in most cathode-ray tubes or, as seems preferable, magnetically by suitable coils located

FIG. 6-9a. Kinescope used in the reproduction of television signals. (RCA.)

just outside the tube envelope. The current passing through the deflecting coils is varied in such a manner as to cover the entire picture area thirty times per second, a process known as *scanning*.[1]

The iconoscope (Fig. 6-10) is the pickup tube for television trans-

[1] There are a number of published papers describing methods of scanning. For example, see V. K. Zworykin, Television, *J. Franklin Inst.*, 217, p. 1, January, 1934; R. D. Kell, A. V. Bedford, and M. A. Trainer, Scanning Sequence and Repetition Rate of Television Images, *Proc. IRE*, 24, p. 559, April, 1936.

mission.[1] As in the cathode-ray tube, a gun (Fig. 6-11) projects an electron stream which is deflected by electrostatic or electro-

FIG. 6-9b. A 10-in. kinescope. (*RCA.*)

magnetic means (usually the latter). However, instead of having a fluorescent screen on the end of the tube, a photoelectric surface

[1] For further information see V. K. Zworykin, The Iconoscope: A Modern Version of the Electric Eye, *Proc. IRE*, **22**, p. 16, January, 1934; R. B. Janes and W. H. Hickok, Recent Improvements in the Design and Characteristics of the Iconoscope, *Proc. IRE*, **27**, p. 535, September, 1939; Harley Iams, G. A. Morton, and V. K. Zworykin, The Image Iconoscope, *Proc. IRE*, **27**, p. 541, September, 1939.

or target is mounted in the large part of the bulb on which is focused the scene or picture to be transmitted. This photoelectric surface is made up of a large number of minute photoelectric cells each of which constitutes one plate of a condenser. The back of the target

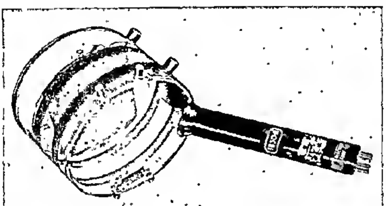


FIG. 6-10. An iconoscope used in the transmitting of television signals (*RCA*.)

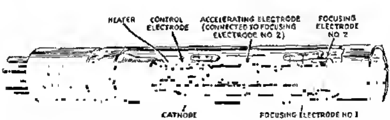


FIG. 6-11. Cutaway view of the electron gun used in a typical iconoscope. (*RCA*.)

is faced with a metal plate which serves as the other plate of the condensers.

Figure 6-12 illustrates the operation of a very small section of the surface. Light striking each cell causes emission and thus produces a charge proportional to the intensity of the impinging light. The electron beam is swept or scanned over the surface of the target, and when it falls on a given cell, the electrons of the beam discharge

the condenser, allowing a current to flow through the resistor R which is proportional to the intensity of the light striking the portion of the target being scanned. The voltage thus produced across the resistor is then amplified and sent out over the television transmitter.

The orthicon represents an improvement in pickup tubes.[1] The picture to be transmitted is focused on a semitransparent photocathode and electrons are emitted from the opposite side. The charges thus built up are scanned by a low-velocity beam to avoid

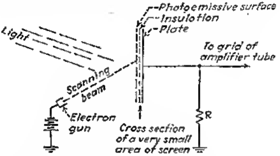


FIG. 6-12. Illustrating the operation of a small section of the surface of an iconoscope.

the secondary emission which causes reduced picture quality at low light levels in the iconoscope. A magnetic field is added with flux lines parallel to the axis of the tube to prevent the electron scanning beam from spreading out, as it would otherwise do at the low velocities used.

A more recent development in pickup tubes is the image orthicon,[2] Fig. 6-13, which differs from the orthicon in three respects: (1) the charge pattern on the target is formed by secondary emission rather than by photoemission, (2) a two-sided target is used, and (3) an electron multiplier is incorporated into the tube. A semitransparent photocathode is used, as in the orthicon, but the electrons emitted from this surface are attracted to another plate

[1] Albert Rose and Harley Iams, The Orthicon, A Television Pickup Tube, *RCA Rev.*, 4, p. 186, October, 1939; The Orthicon, *Electronics*, 12, p. 11, July, 1939.

[2] Albert Rose, Paul K. Weimer, Harold Law, The Image Orthicon—A Sensitive Television Pickup Tube, *Proc. IRE*, 34, p. 424, July, 1946.

which serves as the target (see Fig. 6-14). Here they produce secondary emission and, as the number of secondary electrons exceeds the number of electrons coming from the photocathode, a positive-charge pattern is formed on the target. The low-velocity scanning beam strikes the opposite side of the target where a number of electrons is absorbed from the beam in proportion to the charge on the portion of the target being scanned. The remaining electrons in the beam are reflected back toward the cathode where they are focused magnetically on the surface of the electron gun. The secondary emission produced by impact of these electrons constitutes the first stage of what is usually about a five-stage electron multiplier. The output of this multiplier is then applied to the usual video amplifiers.

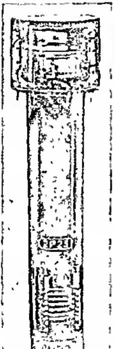


FIG. 6-14.

The three types of television pickup tubes—iconoscope, orthicon, and image orthicon—represent increasing degrees of sensitivity. Thus the iconoscope is capable of producing very satisfactory pictures when the light level is relatively high, the orthicon extends the range into medium light levels, whereas the image orthicon extends the range into lower light levels by a factor of approximately 100 over the orthicon.

X-ray Tubes.[1] When a high-speed electron collides with a body, some of the electrons which constitute that body are set into rapid, vibratory motion. If the impact is sufficiently severe, the disturbed electrons will send off very high frequency radiations known as *Roentgen rays* or *X rays*. These rays have the property of penetration to

[1] There have been a number of engineering applications of X rays, e.g., in inspection of steel castings, but the study of X rays is nevertheless more closely related to the fields of physics and medicine than to engineering, therefore, only a very brief discussion of X-ray tubes is included here. For a general discussion of various types of X-ray tubes and some of their applications see Zed S. Atlee, Design and Application of X-ray Tubes, *Electronics*, **13**, p. 26, October, 1940.

a very high degree and are used, among other things, for making observations through solids. Examples of such use are found in medical and dental work where diseased or broken parts of the human body may be observed and treated and also in industrial applications where castings and metallic joints are examined for flaws, foreign metallic bodies are detected in foods, shoes are fitted to feet, etc.

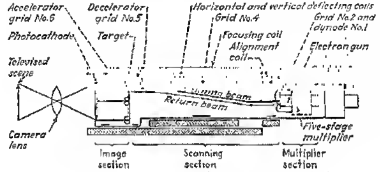


FIG. 6-14. Structural details of image orthicon. (*RCA*.)

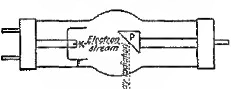


FIG. 6-15. Coolidge X-ray tube.

A sketch of a simple X-ray tube is shown in Fig. 6-15 in which P is the anode target and K is a hot cathode, contained in a highly evacuated vessel. The metallic ring F around the cathode is a focusing shield designed to focus the electron stream on the anode. The face of the anode is inclined at an angle in order to throw the X-ray radiation out the side of the glass bulb. To operate the tube, the cathode is heated as in any hot-cathode vacuum tube, and a potential of a few hundred thousand up to several million volts, depending upon the size of the tube and the nature of the work that it is to perform, is applied between the anode and cathode. The intensity of the X-ray radiations is controlled by

the number of electrons striking the anode target, which in turn is a function of the temperature of the cathode. The plate potential may be varied to control the frequency of the radiations.

X-ray tubes are built in sizes ranging from those which may be held in the hand to those which must be built in sections because of their great length. The latter are used in radiotherapy work and in experimental physics The smaller sizes are used in dental and medical work where visual inspection of the body is desired, and in similar inspection of manufactured articles.

PART II

APPLICATIONS AND CIRCUITS

INTRODUCTION TO PART II

As explained in the Introduction to Part I and in the preface to the second edition, the purpose of Part II is to present basic applications and circuits for vacuum tubes. Chapters 7 and 8 contain material of primary interest to the power engineer, where the term *power engineer* is intended to include all electrical engineers not specifically interested in communication work. Chapters 9 to 14 are of primary interest to the communication engineer. It should not be inferred from this statement, however, that none of the material in these six chapters is of any interest to the power engineer or that that in the first two is of no interest to communication engineers. On the contrary, many applications in the power field, for example, require the use of vacuum-tube amplifiers and even of oscillators and modulators, whereas rectifiers are used extensively in supplying d-c power to vacuum tubes used in communication circuits. The arrangement of the material, however, permits those interested in power and industrial applications to devote the major portion of their time to Chaps. 7 and 8, whereas those interested in communication work may omit Chap. 8 entirely and include only as much of Chap. 7 as seems desirable.

CHAPTER 7

RECTIFIERS

A characteristic of the vacuum tube of prime interest and value to the power, or industrial, engineer is its rectifying property, *i e.*, its ability to convert alternating current to direct current. This property is inherent in a vacuum tube because of its ability to pass electrons from cathode to anode but not in the reverse direction.

An ideal rectifier is one that presents zero impedance to the flow of current in one direction and infinite impedance in the opposite direction. Such a rectifier is equivalent to a switch, closed when one polarity of emf is applied and open under the reverse polarity. In this respect the vacuum tube is not ideal, but the impedance of a gas-filled rectifier tube is extremely low while conducting and virtually infinite during the negative half cycle Thus gas-filled tubes are a reasonably close approximation to a closed switch under positive polarity (plate to cathode) and a virtual equivalent of an open switch under negative polarity. High-vacuum tubes fall much shorter of the ideal conditions under positive polarity and are, therefore, not widely used as rectifiers except where the current (and, therefore, the tube drop) is low The principal application of high-vacuum tube rectifiers is in radio receivers where the load current is so low that the drop in modern high-vacuum tubes is little if any more than in the corresponding gas-filled tubes.

Single-phase, Half-wave Rectifier. The simplest rectifier circuit for use with vacuum tubes is the single-phase, half-wave circuit of Fig 7·1. Winding 1 on the secondary of the transformer supplies the alternating current that is to be rectified, the potential of this winding being such as to produce the desired direct potential at the output. The filament of the rectifier tube is commonly heated by a separate secondary winding 2. The d-c output is then taken off between the filament of the tube and the other side of winding 1 as shown.

The current flow through this type of rectifier when supplying

a resistance load is shown in Fig. 7-2. It is seen to be unidirectional but pulsating. During the half cycle of positive plate potential, current flows in an amount determined by the instantaneous alternating voltage of winding 1, the resistance of the load, and the drop in the tube. As the tube drop is usually low compared with the voltage across the load, the current pulses are, for all practical purposes, half sine waves when a sine wave of emf is impressed. The current flow is, of course, zero during the negative half cycle.

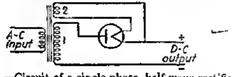


FIG. 7-1.—Circuit of a single-phase, half-wave rectifier.

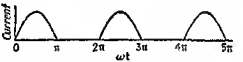


FIG. 7-2.—Wave shape of the current flowing in the output circuit of Fig. 7-1 with resistive load.

The circuit of Fig. 7-1 is seldom used in practice except where a potential is required with little or no current flow (as for grid bias on vacuum tubes, page 422). Such extreme variations in the output current, as indicated in Fig. 7-2, are generally objectionable. These may be smoothed out by the use of filters (see page 182), but the cost of filtering the output of this circuit is excessive if more than a few milliamperes of load current are to be drawn.

Another objection to the use of this circuit is the tendency of the transformer core to saturate. Inspection of Fig. 7-2 shows that the ampere turns in winding 1 of the transformer of Fig. 7-1 will set up flux in the core in one direction only and so tend to cause operation of the transformer above the knee of the saturation curve. Transformer design for such service must be unusually liberal, else the charging current and therefore the losses will be excessive.

Single-phase, Full-wave Rectifier. The most common rectifier circuit for use with single-phase supply is the full wave (Fig. 7-3).

Two tubes are used to rectify both halves of a single-phase supply, so that one tube passes current throughout one half cycle of the supply while the other passes current throughout the next. This process is repeated during each successive cycle, producing an output current (with resistive load) having the wave shape shown in Fig. 7-4. For direct voltages under 500 both tubes are usually incorporated in a single bulb such as the 80, Fig. 3-7, or the 5Z4, Fig. 7-5, using the circuit of Fig. 7-6.

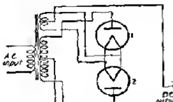


FIG. 7-3.—Circuit of a single-phase, full-wave rectifier.

The positive output terminal in Figs. 7-3 and 7-6 is supplied from a mid-tap on the filament winding of the transformer, whereas

FIG. 7-4.—Wave shape of the current flowing in the output circuit of Fig. 7-3 with resistive load.


FIG. 7-5.—Small, double-anode diode.

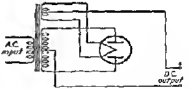


FIG. 7-6.—Circuit of a full-wave, single-phase rectifier, using a single tube with two anodes.

in Fig. 7-1 it was supplied from one side of the filament. In most cases there is little to choose between the two connections. The method of Fig. 7-1 does introduce an additional alternating voltage in the output, equal to half the voltage of the filament winding, but

compared to the other alternating voltages resulting from the rectifying action its effect is negligible.[1] Where tubes with indirectly heated cathodes are used, the connection is made to the cathode.

The two principal objections to the use of the single-phase, half-wave rectifier of Fig. 7-1 are largely avoided in the full-wave rectifiers of Figs. 7-3 and 7-6. The current, though pulsating, drops to zero for only an instant at the end of each half cycle. Secondly, d-c saturation of the transformer core is entirely eliminated, provided the two halves of the transformer produce equal voltages and the tubes have identical characteristics. The current in each tube will then produce equal ampere turns during alternate half cycles, that of one tube setting up flux in one direction and that of the second tube setting up flux in the reverse direction around the

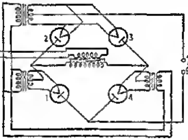


FIG. 7-7.—Single-phase, full-wave, bridge rectifier circuit.

transformer core. Thus a true alternating flux will exist in the transformer.

Single-phase, Bridge Rectifier. Occasionally a bridge circuit (Fig. 7-7) is used. Four tubes are connected in a diamond with only a single secondary winding. During the half cycle that the left terminal of the secondary winding is positive, current flows through tube 2, through the d-c load circuit, and back through tube 4 to the right terminal of the secondary. It will be observed that the plates of tubes 1 and 3 are negative with respect to their cathodes and cannot, therefore, pass current. During the next half cycle the right terminal of the secondary will be positive, and current will flow from there through tube 3, then through the load

[1] See Table 7-1, p. 195, for the magnitudes of the alternating voltages appearing in the output of a rectifier using various connections.

circuit in the same direction as before, through tube 1, and back to the secondary winding. Tubes 2 and 4 are now inoperative owing to their plates being more negative than their cathodes.

The chief advantages of this circuit over that of Fig 7-3 are that (1) higher voltages may be rectified owing to the use of two tubes in series and (2) the transformer secondary winding requires but half as many turns. The principal disadvantages are (1) higher cost due to the need for three filament windings and four tubes as against the one winding and two tubes of Fig. 7-3 and (2) poorer voltage regulation, especially at low output voltages, resulting from the use of two rectifier tubes in series. The higher cost under (1) above is somewhat offset by the decreased cost of the power transformer. In low-voltage, low-power rectifiers this latter saving is almost negligible, so that the bridge circuit is seldom used in any but high-voltage rectifiers. The use of gas-filled tubes in this circuit is especially desirable, since their low internal drop results in reasonably good voltage regulation even with two tubes in series, if the output voltage is not too low.

Polyphase Rectifier Circuits. A three-phase power supply is always used for higher power rectifiers to increase the efficiency, decrease the alternating components in the output, and utilize the vacuum tubes more effectively. Some of the more common polyphase circuits are shown in Figs. 7-8, 7-10, 7-11, 7-13, and 7-16. In these figures all transformer windings that are drawn parallel to one another represent windings on the same transformer and therefore associated with a given phase. In Fig. 7-11, for example, there are one primary winding and two secondary windings per transformer or phase. The tube filaments are generally supplied by a separate transformer which may be energized from any one of the three phases. The secondary of the filament transformer must normally be insulated from its primary for the full output voltage of the rectifier, since it is commonly connected directly to the positive terminal of the rectifier, whereas the negative terminal is grounded. In high-voltage rectifiers a specially designed transformer, or transformers, are therefore required.

Three-phase, Half-wave Rectifier. The simplest polyphase rectifier circuit is that of the three-phase, half-wave rectifier (Fig. 7-8), the operation of which may best be studied with the aid of Fig. 7-9. The sine waves in this latter figure represent the voltages existing between the outer ends of each of the three secondary

windings and the grounded center point of the Y of Fig. 7-8. Inspection of Fig. 7-8 reveals that the center point of the Y is also the negative d-c terminal, whereas the outer terminal of each secondary winding is connected to the plate of a tube. Therefore,

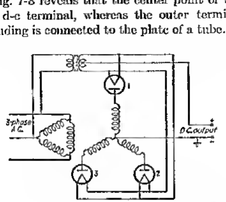


FIG. 7-8.—Circuit of a three-phase, half-wave rectifier.

the horizontal line *oo* of Fig. 7-9 represents the potential of the negative d-c terminal (this being the reference potential), and the sine waves 1, 2, and 3 each represent the potential of the plate of one of the tubes.

The cathode potential of each tube may now be determined, if the tube drop is neglected.[1] With negligible tube drop it is obvious that the plate potential of each tube can never be more positive than its cathode, since such difference in potential would constitute tube drop. On the other hand, the plate of any tube may be more negative than its cathode, since no current will flow with negative plate voltage. Furthermore, it may be seen from Fig. 7-8 that the cathode potential of all three tubes must be the same, since they are all tied together. Therefore, the cathode potential of all three

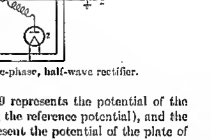


FIG. 7-9.—Illustrating the operation of the circuit of Fig. 7-8 with resistive load.

[1] In most well-designed rectifiers this drop is only a very small percentage of the direct output voltage, and zero tube drop may therefore be assumed without appreciable error. Methods are presented later for correcting for this drop in the design of rectifiers.

tubes at any instant must be equal to that of the most positive anode and must follow the upper envelope of Fig. 7-9, indicated by the solid line *abcde*. . . . If this potential followed any of the other curves shown, one of the anodes would be more positive than its cathode, an impossibility under the assumption of no tube drop. It further follows that tube 1, and only tube 1, will conduct during the interval t_1t_2, since its cathode and anode are at the same potential, whereas the anodes of the other two tubes are more negative than their cathodes

At time t_2 the voltage of secondary winding 2 will just equal that of winding 1, and both tubes 1 and 2 will tend to pass current An instant later, however, the voltage of winding 2 will exceed that of 1, and tube 1 will have a negative voltage applied between plate and cathode. It will then cease to conduct, and tube 2 will pass current until time t_3 Similarly tube 3 will pass current from time t_3 to t_1', while tubes 1 and 2 are inactive. The process will repeat itself during each ensuing cycle with the conducting period of each tube as indicated. The circuit acts very much as a commutator, switching the load current from one secondary winding to the next every third of a cycle

The instantaneous output voltage is evidently given by the distance between the upper envelope and the line *oo*. This follows immediately, since the two output terminals are connected directly to the cathodes of the tubes and to the center point of the Y, respectively. Evidently the output voltage pulsates between a maximum and a minimum three times per cycle, and if the load is resistive the current will pulsate in the same manner.[1] Evidently, however, the magnitude of these pulsations is less than for the single-phase, full-wave rectifier, since the current never goes to zero

The circuit of Fig. 7-8 has the very apparent disadvantage of a residual d-c flux in the core of each transformer as in the circuit of Fig. 7-1. This may be avoided by using the distributed-Y circuit of Fig. 7-10, which operates in all ways exactly as does that of Fig. 7-8, although each secondary has two windings supplying voltage to two different tubes The d-c components in each winding of a given transformer are flowing in opposite directions so that the

[1] In the first seventeen pages of this chapter the term *load* will be construed to include any filters used. Evidently, then, a resistive load means no filtering, since filters are made up of condensers and inductances.

residual d-c flux is zero. This may permit a cheaper design of transformer, even though the total number of turns required to produce the same direct voltage is greater than for the circuit of Fig. 7-8.

Three-phase, Full-wave Rectifier. The full-wave circuit of Fig. 7-11 uses the same type of power transformer as that of Fig. 7-10, but all six secondary windings have one terminal connected to the

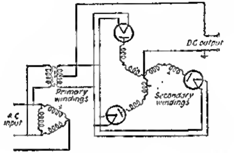


FIG. 7-10.—Circuit of the distributed-Y type of three-phase, half-wave rectifier.

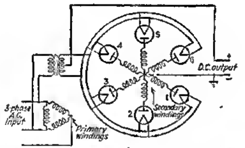


FIG. 7-11.—Circuit of a three-phase, full-wave rectifier. (Sometimes called a six-phase rectifier.)

center point of the Y, and each supplies voltage to a separate tube. Very little gain is realized in using this circuit over that of the three-phase distributed Y (Fig. 7-10) except in the lower ripple obtained, since each tube must still pass the full direct current during its conducting period of one-sixth of a cycle, and the direct voltage, with given transformers, will be lower. The curve of output voltage is shown by the heavy line in Fig. 7-12, the conduction period of each tube being indicated.

If the secondary windings of the full-wave circuit are split into two three-phase, half-wave rectifiers and the mid-points of the two Y's are connected together through a mid-tapped inductance or *balance coil* (also known as *interphase reactor*), Fig. 7-13, each tube will conduct current for *one-third* of a cycle and will, therefore, for a given load, have a peak current only half as great as required in the polyphase circuits discussed thus far. The presence of the

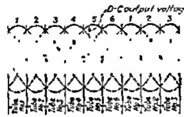


FIG. 7-12.—Illustrating the operation of the circuit of Fig. 7-11 with resistive load.

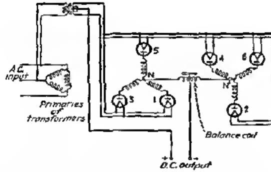


FIG. 7-13.—Three-phase, double-Y rectifier circuit.

balance coil has the effect of making each of the two Y's act as a separate rectifier by absorbing the instantaneous difference in voltage between the two, thus causing each tube to carry current as though operating in the three-phase, half-wave circuit of Fig. 7-8.

The currents and voltages are shown in Fig. 7-14. In this figure the upper (solid) envelope of (*a*) and (*b*) again represent the potential of the cathodes of all tubes, and the sine waves 1, 2 . . . 6 repre-

sent the potentials of each of the six anodes *but with reference to the center point of the particular Y with which the tube is associated,* not with reference to the negative d-c terminal. Therefore, the

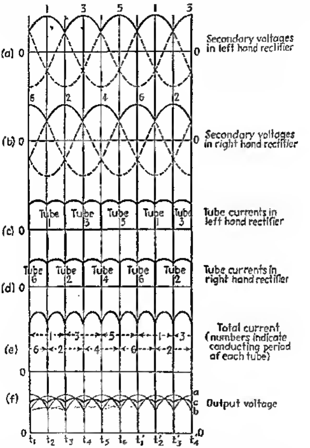


FIG. 7-14.—Illustrating the operation of the circuit of Fig. 7-13 with resistive load.

line oo in (a) represents a different reference potential from the one represented by oo in (b), the former representing that of point N (Fig. 7-13), whereas the latter represents that of N'. If the negative d-c terminal were used as reference, the potential of the

center points N and N' of the two Y's would be found to vary up and down, above and below this reference potential, by the drop across each half of the balance coil, the center point of one Y being as much above the potential of the d-c terminal as that of the other was below. The anodes of tubes 1 and 6 are, as a result of this action, more positive than the negative d-c terminal by *exactly the same amount* during the interval t_1t_2, the voltage of the balance coil adding to the voltage of transformer winding 1 and subtracting from that of 6 during the first half of this period and performing the reverse function during the second half to provide the same total voltage on the anode of each tube. Similarly, the anodes of tubes 1 and 2 are equally positive during the interval t_2t_3, etc., and it is quite evident from the figure that each tube conducts for one third cycle, covering two pulses of the total output current (Fig. 7-14*e*, which is the current flowing with resistance load and negligible tube drop).

In (*f*) of Fig. 7-14 the voltage curves of (*a*) and (*b*) are reproduced as *a* and *b*, superimposed one on the other. The difference between these two curves is at every instant equal to the difference in the voltages of the left-hand and right-hand rectifiers, which is the voltage across the balance coil. Since the output is taken from a midtap on this coil, the output voltage must follow a curve which is the instantaneous average of curves *a* and *b*, as curve *c*. It is also evident from these curves that the voltage across the balance coil is far from sinusoidal, with a peak voltage equal to the difference between the maximum and minimum voltages of curves *a* and *b*.

Three-phase Bridge Rectifier. The three-phase, double-Y circuit of Fig. 7-13 permits two three-phase, half-wave rectifiers to operate in parallel. Two half-wave rectifiers may also be operated in series as indicated in Fig. 7-15*a*, where the filament supply and the primary windings of the power-supply transformers are omitted for simplicity. The voltage and power output of such a combination are double that of a single rectifier, but the current capacity is the same.

If the tubes of the lower rectifier are reversed in polarity, as in *b* of the same figure, the output power and voltage will not be changed if the center points of the two Y's are tied together as shown to make the voltages of the two rectifiers additive. It may now be seen that points A and A' are at exactly the same potential, the

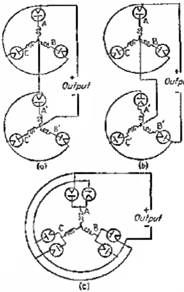


FIG. 7-15. Two three-phase, half-wave rectifiers operated in series are the substantial equivalent of the three-phase, full-wave bridge circuit.

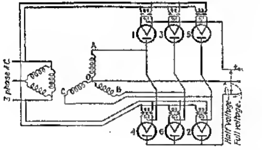


FIG. 7-16.—Full-wave, three-phase bridge rectifier circuit.

same thing being true of points B and B' and points C and C'. It is, therefore, possible to connect the three lower tubes directly to points A, B, and C, respectively, and dispose of one set of transformer windings entirely, as in c of Fig. 7-15.

The circuit of Fig. 7-15c is known as a three-phase, full-wave

bridge circuit and is shown in complete form in Fig 7-16. As indicated in the preceding paragraphs, two tubes operate in series, one to provide a path from the transformers to the positive d-c terminal, and the other to provide a return path from the negative terminal. Six filament transformers are shown, although the cathodes of tubes 1, 3, and 5 could all be heated from a single transformer of three times the rating of those shown. The use of separate transformers is usually preferable, since it is then necessary to keep but a single spare for emergency use.

The curves of transformer secondary voltages for this circuit are shown by the sine waves in Fig. 7-17 with the center of the Y as reference point. Evidently the sine wave OA also represents the potential of the anode of tube 1 and the cathode of tube 4, OB serves the same function for tubes 3 and 6, and OC serves for tubes

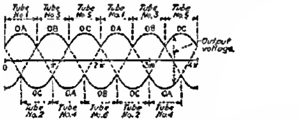


Fig. 7-17.—Illustrating the wave shape of the output voltage of the rectifier circuit of Fig 7-16. Center point of Y used as the reference point.

5 and 2. Furthermore, the cathodes of tubes 1, 3, and 5 are all tied together, and since under the assumption of no tube drop a cathode cannot be more negative than its anode, the potential of these cathodes must follow the upper envelope as indicated by the upper heavy curve. A similar analysis shows that the potential of the anodes of tubes 2, 4, and 6 must follow the lower envelope. Since the cathodes of tubes 1, 3, and 5 are tied to the positive d-c terminal while the anodes of tubes 2, 4, and 6 are tied to the negative d-c terminal, it is obvious that the direct output voltage is given by the distance between the two envelopes as indicated by an arrow in Fig. 7-17.

The conduction period of each tube may now be determined from the curves The anode potential of tube 1 follows curve OA; the cathode potential follows the upper envelope. Therefore,

tube 1 must be conducting during the time these two curves coincide and idle at all times when curve OA is more negative than the upper envelope. Applying this analysis to all six tubes shows that each conducts current for one-third cycle throughout the periods indicated in Fig. 7-17.

The curves of Fig. 7-17 may be replotted as in Fig. 7-18 using the negative d-c terminal as reference point and plotting the total Y voltages AB, BC, and CA instead of the phase voltages as in Fig. 7-17. In this case the instantaneous output voltage is given by the vertical distance between the heavy envelope line and the X axis. This figure more clearly portrays the wave shape of the output voltage, but its construction is not readily followed unless obtained from Fig. 7-17.

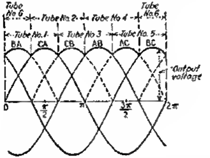


FIG. 7-18.—Illustrating the wave shape of the output voltage of the rectifier circuit of Fig. 7-16. Negative terminal used as the reference point.

A tap taken off the center point of the Y-connected secondaries is shown in Fig. 7-16, giving a voltage to ground (negative lead) only half that obtained between outside terminals. As may be seen from Fig. 7-15b, the power at this half-voltage tap is supplied by a half-wave, three-phase rectifier using only half the tubes of Fig. 7-16. This often provides a cheap and effective means of securing a lower voltage tap where the power demand is small compared to that taken off the outside leads. If too heavy a load is taken off this center lead, saturation of the transformer cores is likely to result from the unbalanced direct component present.[1] Furthermore, tubes 2, 4, and 6 will be more heavily loaded than

[1] This may be avoided by using six secondary windings in the distributed-Y construction of Fig. 7-10.

the other three, and the capacity of the rectifier at the higher voltage may be unduly reduced.

Choice of Tubes and Circuits. The various types of thermionic tubes used in rectifier service were described in Chaps 3 and 4. These were (1) high-vacuum, (2) mercury-vapor, (3) mercury-arc, (4) ignitron, and (5) tungar. For rectifier service these tubes are generally diodes, although gas-filled triodes are often used in rectifier service where it is desired to have the output voltage variable at the will of the operator (see page 237)

High-vacuum tubes are used in rectifier service only when the voltage is so high that gas-filled tubes are likely to arc back, or when the current demand is so low that the drop in a high-vacuum tube is little if any greater than that in a gas-filled tube, as in low-power rectifiers for radio receivers The drop in the high-vacuum tube, when the load current is more than a few hundred milliamperes, is much higher than that in the gas-filled tube. This means poorer regulation and greater tube loss, therefore, lower efficiency. Consequently the use of high-vacuum tubes is to be avoided

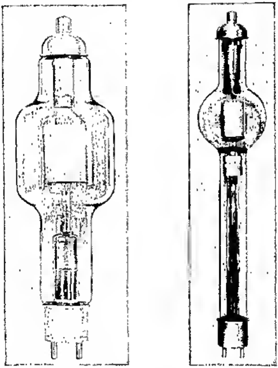
The circuits commonly used for low-power rectifiers using high-vacuum tubes are the single-phase, half-wave circuit of Fig 7-1 for rectifiers designed to deliver a direct voltage with virtually no current, or the full-wave circuits of Fig. 7-3 or 7-6 if a small but appreciable current is to be delivered. Rectifiers designed to supply grid bias for large triode or multigrid tubes are examples of the application of the circuit of Fig. 7-1; rectifiers designed to deliver output voltages not to exceed about 1000 volts and currents not much in excess of perhaps 200 ma fall generally within the second classification.

At higher power outputs and at higher voltages, the three-phase, double-Y circuit of Fig. 7-13 is most commonly used with high-vacuum tubes such as that of Fig. 7-19*a*. This circuit causes two tubes to operate in parallel, thus reducing the tube drop, and provides an output with a minimum ripple or pulsation.

At very high voltages the single-phase bridge circuit is sometimes used to reduce the inverse voltage applied to the tubes on the negative half cycle. Figure 7-20 shows such a rectifier. Since two tubes are in series, each tube receives but half the inverse voltage. On the other hand the total tube drop is increased so that bridge

circuits are to be avoided when high-vacuum tubes are used unless the load current is so low that the tube drop is not excessive.

The tubes used in very high-voltage rectifiers must be especially designed to prevent flashover, either inside or outside the envelope. Note particularly the rather large spacing between anode and



(a) (b)

FIG. 7-19.—High-vacuum rectifier tubes. The tube shown in (a) is designed for operation at reasonably high values of current and voltage, while that in (b) is designed for use at very high voltages.

cathode terminals (at opposite ends of the tube) in the tubes shown in Fig. 7-20. The tube of Fig. 7-19b is designed for similar service but with even greater separation of its terminals.

Some form of gas-filled tube is used in most high-power rectifiers. The higher voltage units, from perhaps 2000 or 3000 to about

20,000 volts, use either the hot-cathode, mercury-vapor tube, or the ignitron, the latter being preferred for rectifiers of higher power output; yet both these tubes are often used at lower voltages. Mercury-arc rectifiers have been widely used in sizes as large as 3000 kw at voltages of 600 and 1200 volts in railway service, but ignitrons are now displacing them; thus the mercury-arc rectifier

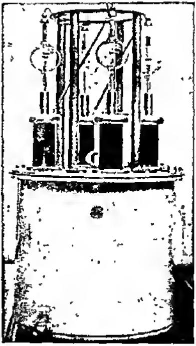


FIG. 7-20.—100,000 volt, 0.5 amp., single-phase rectifier using the full-wave bridge circuit. (*General Electric Co.*)

is rapidly losing its importance. The tungar tube is used at voltages below about 100 and for currents of 1 to 15 amp. It is especially useful in battery charging.

Mercury-vapor tubes are most commonly used in the full-wave, single-phase circuits of Figs. 7-3 and 7-6 for low-power outputs, although they may be used in the bridge circuit of Fig. 7-7 if the

voltage is high. The advantage of this latter circuit for high voltages lies in the use of two tubes in series, reducing the inverse voltage per tube on the negative half cycle. Since mercury-vapor tubes are much more susceptible to arc back than are high-vacuum tubes, any reduction in inverse voltage is highly desirable.

The three-phase, bridge circuit of Fig. 7-16 is almost invariably used with mercury-vapor tubes at high-power outputs. Being a bridge circuit, the inverse voltage on the tubes is lower than in any other three-phase circuit. Operation of two tubes in series does, of course, tend to give poorer regulation, but the drop in mercury-vapor tubes is so low (about 15 volts) that even with two in series it is nearly negligible in a high-voltage rectifier.

Ignitrons may be used in the bridge circuit of Fig. 7-16. However, it is frequently desirable that the cathodes of all tubes be kept at the same potential to simplify construction problems, and the circuits of Figs. 7-11 and 7-13 are then used. Since the inverse voltage of these circuits is higher than that of Fig. 7-16, the maximum safe output voltage is less than when the bridge circuit is used.

Mercury-arc tubes are built with but one cathode for all anodes and therefore cannot be used where the cathodes must be operated at different potentials, as in the bridge circuits. The circuits of Figs. 7-11 and 7-13 are, therefore, most commonly used. Mercury-arc rectifiers cannot be designed to operate at very high voltages, generally not more than a few thousand. This rather low voltage limit is due in part to the low inverse voltage rating of the mercury-arc tube, which is a result of placing all anodes in a single chamber (see page 107). Thus the mercury-arc tube has a greater tendency to arc back than either the mercury-vapor tube or the ignitron.

Tungar tubes are used in the single-phase, half-wave circuit of Fig. 7-1 or in the full-wave circuit of Fig. 7-3. They are not built with two anodes in a single bulb, so the circuit of Fig. 7-6 is not used. (See also Fig. 4-19.)

Filters. For some purposes the output current of a three- or six-phase rectifier may be sufficiently constant, but in a large majority of cases a much higher degree of constancy is required. By the use of filters almost any degree of constancy or smoothness may be achieved.

Two principal types of filters are used with rectifiers. The condenser-input, or pi-section, filter is illustrated in Fig. 7-21; the inductance-input, or L-section, filter is shown in Fig. 7-22. In the

pi-section filter the condenser C_0 charges during the time that a tube is conducting, and during any nonconducting period it supplies the full load current through the inductance L, with a drop in voltage determined by the size of the condenser and the length of the nonconducting period. The inductance tends to maintain the flow of current to the load at a constant value, the small variations that it permits being largely smoothed out by the condenser C. The L-section filter operates in a similar manner except that the inductance, if operating properly, will demand a continuous flow of tube current at all times during the cycle. This type of filter is therefore not ordinarily used with a single-phase, half-wave recti-

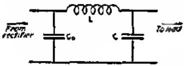


FIG 7-21.—Circuit of a pi-section (condenser-input) filter.

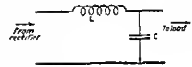


FIG. 7-22.—Circuit of an L-section (inductance-input) filter.

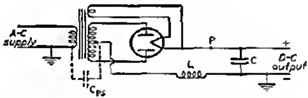


FIG. 7-23.—Circuit indicating the reduction in filtering which may take place with the choke placed in the negative lead.

fier where no tube is available to pass current to the filter during the negative half cycle of the supply voltage.

The inductance, or choke, of a filter is commonly located in the positive rather than the negative lead. This is to prevent by-passing of the higher ripple frequencies around the choke through capacitances to ground.[1] Figure 7-23 illustrates this phenomenon for a single-phase, full-wave rectifier with an L-section filter. The rectifier is very commonly grounded at the negative terminal as shown, and the primary wiring is usually grounded or at least has

[1] See F. E. Terman and S. B. Pickles, Note on a Cause of Residual Hum in Rectifier-filter Systems, *Proc. IRE*, **22**, p. 1040, August, 1934.

a large capacitance to ground. The capacitance C_{ps} between primary and secondary windings then completes the by-pass circuit when the choke is placed in the negative lead as shown. The capacitance C_{ps} will evidently pass the higher frequency components more readily than the lower, and these components, being well up in the audio spectrum, will cause serious disturbance in any communication equipment being supplied by this rectifier and may cause undesirable reactions in many industrial applications.

Some power transformers, especially those used in radio receivers, are supplied with an electrostatic shield between primary and secondary windings to prevent disturbances in the power line from passing through the capacitance between the two windings. Such a shield consists of a conducting cylinder wrapped around the core between the two windings but split along an element of the cylinder to prevent its acting as a short-circuited turn. It is grounded when the transformer is installed. In such a trans-

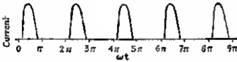


FIG. 7-24.—Current flowing in the circuit of Fig. 7-1 when used with a pi-section filter.

former the capacitance C_{ps} of Fig. 7-23 is eliminated, but the capacitance between the secondary and the electrostatic shield is an even more effective by-pass around the choke.

Where a high degree of filtering is not required, the choke may be placed in the negative lead with a possible saving in insulation costs for high-voltage units. The voltage to ground from the choke in the circuit of Fig. 7-23 is only that of the choke itself, whereas it will be equal to the direct output voltage plus the choke voltage if the choke is placed in the positive lead, at point P.

Current Flow in Rectifiers When Filtered. Figures 7-24 and 7-25 show the currents in the half- and full-wave single-phase rectifiers supplying a pi-section filter. The current flow through the tubes is seen to be quite different in nature from that of the rectifier operating alone (Figs. 7-2 and 7-4), in that the current flows in a series of short, high-peaked pulses. This is characteristic of pi-section filters, since the condenser C_o acts almost as a

short circuit across the rectifier terminals. Because of this extreme demand upon the rectifier tubes the pi-section filter is seldom used on any but very low-power rectifiers, as in radio receivers, where it is often preferred to the L-section filter because of its somewhat higher output voltage and lower cost for a given amount of filtering.

The currents in a full-wave, single-phase rectifier operating into an L-section filter are shown in Fig. 7-26. Here the tube current is nearly constant throughout the conducting period, with one tube picking up the load current as the other drops it. The peak value of current flowing through a tube for a given average value of load current is obviously much lower than for the other type of

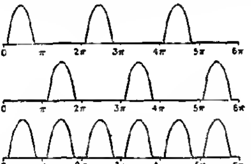


FIG. 7-25.—Currents flowing in the circuit of Fig. 7-3 when used with a pi-section filter. The two upper curves show the current in the individual tubes, while the bottom curve shows the current in the common lead.

filter. For this reason L-section filters are always used with high-power rectifiers.

As stated in a preceding paragraph the L-section filter is not ordinarily used with a half-wave, single-phase rectifier, since such a rectifier cannot pass current throughout a full 360 deg. When it is used, the tube will conduct for a period greater than 180 deg but less than 360 deg as shown in Fig. 7-27. Conduction during the negative half cycle of the supply voltage is due to the induced emf of the inductance exceeding the supply emf and thus maintaining the plate of the tube more positive than its cathode. The filtering action with this circuit is much less effective than with the full-wave circuit, as may be seen by comparing Figs. 7-26 and 7-27.

The addition of a resistance as in Fig. 7-28 will improve the per-

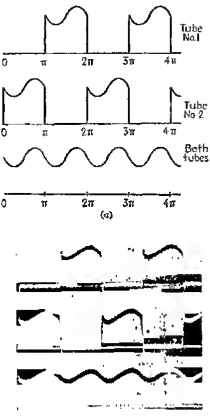


FIG. 7-26.—Current flowing in the circuit of Fig. 7-3 when used with an L-section filter. The two upper curves in (a) show the current in the individual tubes, while the bottom curve shows the current in the common lead. (b) is an actual oscillogram.

formance of a half-wave rectifier with L-section filter.[1] The inductance will then discharge through this resistance during the negative half-cycle of the supply voltage and, if of sufficient size, will maintain a continuous flow of current, as in Fig. 7-29, quite similar to the total current with a full-wave rectifier as shown in

[1] See M. B. Stout, Behavior of Half-wave Rectifiers, *Electronics*, 12, p. 32, September, 1939.

the lowest curve of Fig. 7-26a. Obviously the presence of this additional resistor increases the losses and lowers the efficiency while decreasing the ripple.

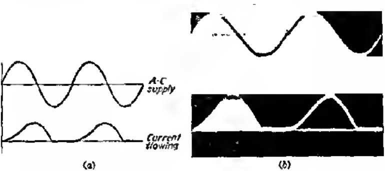


FIG. 7-27.—Current flowing in the circuit of Fig. 7-1 when used with an L-section filter. (b) is an actual oscillogram.

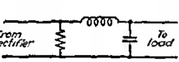


FIG. 7-28.—Illustrating the use of a bleeder resistance to improve the performance of an L-section filter when used with the circuit of Fig. 7-1.

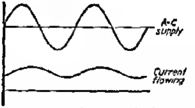


FIG. 7-29.—Current flowing in the inductance of Fig. 7-28, when used with the rectifier of Fig. 7-1.

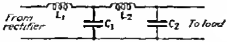


FIG. 7-30.—Circuit of a two-section, inductance-input filter.

In case the simple, single-section filters just described are insufficient to provide the required smoothness of output, additional L-type filter sections may be added as in Fig. 7-30. It is seldom necessary to add more than one additional section.

Tuned-circuit Filters. It is possible to incorporate tuned circuits into the filter to provide increased filtering action at one or

more frequencies. These may consist of either series-resonant units in shunt with the circuit or parallel-resonant units in series, or both may be used in combination. The principal use of tuned circuits is in reducing the fundamental component to a very low level, so low as would require a filter of prohibitive cost with a conventional "brute-force" filter of the pi- or L-section type. An objection to the use of tuned circuits is that, while they are more effective than brute-force filters in filtering the frequency to which they are resonant, they are much less effective at other frequencies. Thus a tuned filter, tuned to the principal frequency component

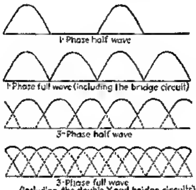


FIG. 7·31.—Wave shape of the output voltage of various rectifier circuits (no filter).

in the output of the rectifier, may be entirely ineffective at higher harmonics unless supplemented by additional untuned filtering.

L-section Filter Design. An exact design of a rectifier and filter is a very difficult process because of the complex nature of the currents flowing, but for most purposes a design of sufficient accuracy may be obtained by making a few simple assumptions. Let the filter be of the L-section type and let the voltage supplied to the filter be equal to the rectified voltage of the transformer; *i.e.*, the tube drop and the transformer reactance drop are to be neglected. Neither of these two factors is entirely negligible, but their principal effect is to cause a slight reduction in the output voltage which may be compensated for by slightly increasing the secondary voltage of the transformer, as illustrated in the example

following this discussion. The voltage supplied to the filter by the various types of circuits will then be as shown by the solid lines in Fig. 7-31, where the heights of the half sine-wave pulses are the crest values of the transformer secondary voltages.

An L-section filter may now be designed with the assistance of Table 7-1.[1] The alternating current flowing through the inductance may be determined with the aid of Fig. 7-32, which is the equivalent circuit considering only the alternating, or ripple, voltages produced by the rectifier and neglecting transformer-reactance drop, tube drop, and choke resistance. Evidently the current is equal to the impressed voltage divided by the total impedance of the filter and load. However, the impedance of the condenser C

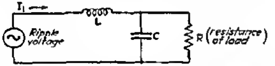


FIG. 7-32.—Equivalent circuit for ripple frequency components of a rectifier.

in a satisfactory filter is only a small fraction of that of the choke L. Thus the condenser and load, being in parallel, have a negligible effect on the magnitude of the alternating current flowing, and it is possible to write

$$I_1 = \frac{E_1}{\omega L} \tag{7-1}$$

where E_1 = alternating voltage applied to filter at ripple frequency $\omega/2\pi$ (obtained from Table 7-1)

I_1 = alternating current flowing through filter at ripple frequency $\omega/2\pi$

L = inductance of coil, henrys

The alternating voltage applied to the filter is anything but a sine wave, as evidenced by the components listed in Table 7-1, so that a complete solution would require that the current flowing at each harmonic frequency be determined; but if the design is such as to keep the current at the *lowest* frequency within satisfactory

[1] The table is on p. 195. Appendix D outlines the methods of determining the numerical values in this table.

limits, all higher harmonics should be negligible. The solution is, therefore, generally carried out at the lowest ripple frequency only.

The alternating voltage appearing across the load terminals may be determined by multiplying the alternating current I_1 by the impedance of C and R in parallel. Actually the condenser reactance is very much smaller than R, so that, for all practical purposes,

$$E_R = \frac{I_1}{\omega C} = \frac{E_1}{\omega^2 LC} \tag{7-2}$$

where E_R = ripple voltage across load
C = capacitance of filter condenser, farads

The known factors in Eq. (7-2) are E_R, E_1, and ω, the first factor being given in the rectifier specifications and the last two being obtained from Table 7-1. The equation may, therefore, be rewritten to solve for the product of L and C:

$$LC \geqq \frac{E_1}{\omega^2 E_R} \tag{7-3}*$$

The $>$ sign is used in this equation, since the specifications normally give the *maximum permissible ripple*. Thus values of L and C giving a product greater than that due to the equality sign in Eq. (7-3) are permissible since this will result in a lower value of ripple voltage E_R than was specified.

Minimum Capacitance for Filter. Evidently an infinite number of combinations of L and C will satisfy Eq. (7-3), and the final choice of a condenser and coil must be governed by economic considerations, to be made only after a study of cost figures. However, there is a minimum value for both L and C below which the filtering action will be unsatisfactory. For C this minimum value is determined by the type of load supplied by the rectifier. Since the series inductance L permits only slow changes in current through the filter, any sudden variations in load demand must be handled, momentarily at least, by the condenser. For example, if the load consists of a vacuum-tube amplifier, there will be continual variations in the current being drawn from the rectifier even though the average current may remain the same. These variations may occur either at a radio or at an audio frequency. If they occur at audio frequencies, the condenser capacitance must be quite large to handle variations in current at a frequency of, say,

30 cycles/sec without permitting appreciable changes in the output voltage of the rectifier. The problem is frequently covered in specifications by stating that the rectifier shall not present more than a given maximum impedance in series with the load at a given minimum frequency. Since the impedance of the condenser is many times lower than that of the inductance, this specification virtually gives the minimum permissible capacitance of the condenser. The relation may be expressed mathematically as

$$C \geqq \frac{1}{\omega_0 Z} \tag{7-4}^*$$

where Z is the maximum permissible impedance of the rectifier at the specified frequency $\omega_0/2\pi$ (This point is fully illustrated in the ensuing example.)

Minimum Inductance for Filter. The purpose of the choke L is to maintain nearly constant current flow through the filter throughout the cycle of the supply voltage. If the load is gradually reduced, a point is reached where the current flowing is so small that the effect of the choke becomes slight; in other words, the inductance is unable to maintain constant current, and the tubes begin to conduct in short pulses as with the pi-section filter The power flow from the transformer is then no longer continuous, a condition that should be avoided.

Since the minimum size of inductance that will avoid the difficulty just referred to is dependent upon the load current flowing, it is necessary to assume some minimum load at which the rectifier is to be operated. Let this current be I_{min}

For the choke to perform properly, the crest value of the alternating current flowing through it must never exceed the direct current. This may be seen by noting that the total instantaneous current flowing through the inductance at any instant is equal to the sum of the direct and alternating components; consequently the total instantaneous current will never fall to zero so long as the crest value of the alternating component is equal to or less than the direct, or

$$\sqrt{2}\, I_1 \leqq I_{min}$$

or, from Eq. (7-1),

$$\frac{\sqrt{2}E_1}{\omega L} \leqq I_{min}$$

Solved for L this equation gives

$$L \geq \frac{\sqrt{2}\,E_1}{\omega I_{min}} \qquad (7\text{-}5)^*$$

where ω is 2π times the ripple frequency. The equality sign in this equation gives the minimum L for satisfactory operation.

Evidently Eqs. (7-3), (7-4), and (7-5) each set a *minimum* relation for the design of the filter constants. Ordinarily the product of the minimum C and the minimum L, as determined by Eqs. (7-4) and (7-5) when using the equality signs, is less than the minimum LC product obtained from Eq. (7-3), so that some leeway is permitted in determining the most economical L and C. If it should so happen that the product of the minimum L and the

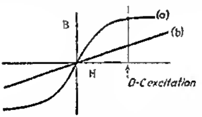


Fig. 7-33.—Illustrating the effect of an air gap on the inductance of a choke. Curve (a) no air gap, (b) with air gap. (Hysteresis neglected.)

minimum C exceeds the minimum LC product obtained from Eq. (7-3), it will of course be necessary to use these values notwithstanding. The result will be a filter capable of providing a lower percentage ripple than that specified.

Once L is determined from the foregoing considerations, its resistance may be estimated and the complete calculations of rectifier performance carried out as outlined in the ensuing sections.

Design of Choke.[1] The inductance of the filter choke L is greatly affected by the magnitude of the direct current flowing through it. If the hysteresis effect is neglected for the moment, the magnetization curve of the iron core of the choke may be shown as in a (Fig. 7-33). Since the inductance of a coil is proportional

[1] The design of a choke for rectifier service is presented in some detail by Reuben Lee, Reactors in D-c Service, *Electronics*, 9, p. 18, September, 1936. For further information see C. R. Hanna, Design of Reactances and Transformers Which Carry Direct Current, *Trans. AIEE*, 36, p. 155, 1927.

to the slope of the B-H curve, it is evident that the maximum average inductance is obtained when no direct current is present and with an alternating current of such magnitude as not to cause the flux to rise above the knee of the curve. If a direct current is present such as to magnetize the core to the point 1, the inductance presented to a small alternating current (as the ripple current in a rectifier) will evidently be proportional to the slope of the B-H curve at that point and will be very much less than that presented with no direct current. The permeability of the core as it affects the alternating current under these conditions is commonly known as the *incremental permeability*, since it is found by superimposing small increments of current on the constant direct current.

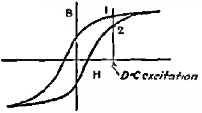


FIG. 7-34.—Illustrating the effect of the hysteresis loop on the inductance of a choke.

Actually the relation between B and H for an iron core is given by the hysteresis loop (Fig. 7-34). Here it is obvious that the effective inductance to the alternating component is dependent not only on the magnitude of the direct component of current but on the previous magnetic history of the iron as well. Thus with the d-c excitation shown, the flux density may be represented by either point 1 or point 2 depending on whether the direct current was decreased or increased to reach the excitation indicated. Furthermore the alternating component itself will tend to cause the magnetization to follow a small hysteresis loop around the point of d-c magnetization, and the effective inductance is proportional to the average slope of this loop. This value is generally smaller than that which would be obtained if there were no hysteresis, as assumed in the preceding paragraph.

The effective inductance of a choke may be increased by inserting an air gap in the magnetic circuit. If hysteresis is again neglected for the sake of simplicity, the magnetization curve will

become as in b (Fig. 7-33). It may be seen that the inductance that such a coil will present to a small alternating current in the presence of a direct component of the magnitude indicated is markedly greater than that of the choke without an air gap. The length of the gap required depends upon the direct current that the choke is to pass and upon the constancy required of the inductance as the direct current is varied. It is generally only a small fraction of the total length of the magnetic circuit, often consisting of a butt joint of the laminations with, perhaps, a thin piece of fish paper or other suitable nonmagnetic material inserted as a spacer.

If the air gap is made very small, the effective inductance will vary somewhat with the direct current, being much larger at low currents than if the gap were of such length as to hold the inductance essentially constant over a wide range of current. Such a unit, known as a *swinging choke*, is commonly used in filters, especially as the first choke in a two-stage filter (L_1 in Fig. 7-30). This type of choke is especially valuable in this location, since it makes possible the satisfying of Eq. (7-5) at much lower values of I_{min} than would be economically possible with a choke containing an air gap of such length as to maintain a nearly constant inductance. The second choke L_2 is of the constant-inductance type which provides adequate filtering at heavy loads where L_1 is small.

Selection of Tube. The ratio of the peak anode tube current to the average or direct rectifier current, given in Table 7-1, is based on the assumption of an infinite inductance and, therefore, a rectangular wave shape of the tube currents. Figure 7-35a shows the current flowing in each tube of a full-wave, single-phase rectifier under these assumptions. Actually, however, the filter inductance is finite, and the rectifier-output current consists of a direct component with several sinusoidal alternating components superimposed. The magnitudes of these alternating components may be determined in the manner already described and the true peak current computed by addition, although for all practical purposes the ripple frequency component is so much greater than any of the others that sufficient accuracy will be realized by considering it alone.

Let that portion of the ripple frequency current flowing through each tube be aI_1, where a represents the fraction of the total ripple current I_1 which flows through one tube. In all the circuits of Table 7-1, except the double-Y, each tube carries the total load

current during its conduction period and a equals one, but in the double-Y two tubes conduct simultaneously and a equals 0.5. Thus a is given by line 1, part C, of the table.

The ripple current per tube, aI_1, is shown in Fig 7-35b for the full-wave, single-phase rectifier Its phase position is found by noting that it must lag the fundamental component of the ripple voltage by virtually 90 deg and that the positive crest of this voltage coincides with the positive crest of the transformer secondary voltage and therefore with the mid-point of the tube-current

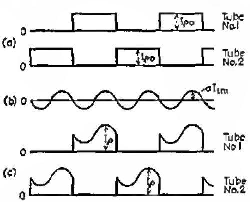


FIG 7-35 - Curves showing the increase in peak current per tube caused by the ripple frequency component in the rectifier

pulse. Summation of the currents in curves a and b produces the two tube currents of Fig 7-35c. We may, therefore, write

$$I_p = I_{p0} + aI_{1m} \tag{7-6}$$

where I_p = actual peak tube current

I_{p0} = peak tube current as computed from Table 7-1.

I_{1m} = crest value of alternating component of current in rectifier output

a = ratio of tube current to rectifier output current (same as ratio of peak anode current to direct current given in line 1, Part C, Table 7-1)

The current I_{1m} is obtained by multiplying the rms current I_1 from Eq. (7-1) by $\sqrt{2}$. Then, if I_0 represents the direct current

in the rectifier output, $I_{p0} = aI_0$ and Eq. (7-6) may be rewritten

$$I_p = a\left(I_0 + \sqrt{2}\,\frac{E_1}{\omega L}\right) \tag{7-7}^*$$

It should be noted that *Eqs.* (7-6) and (7-7) *will not hold true unless the relation of Eq.* (7-5) *is satisfied.* If Eq. (7-5) is not satisfied, the alternating component of current will not be even

TABLE 7-1. RECTIFIER DESIGN DATA FOR TYPICAL RECTIFIER CIRCUITS
(Choke-input filter with infinite inductance assumed)

	Types of circuits					
	1φ full wave Fig. 7-3	1φ bridge Fig. 7-7	3φ half wave Fig. 7-8	3φ distributed Y Fig. 7-10	3φ double Y Fig. 7-13	3φ bridge Fig. 7-10
A. Ratio of alternating voltages to direct output voltage (assuming no losses):*						
1. Transformer secondary, rms (per leg)	1.11†	1.11	0.855	0.494‡	0.855	0.428
2. Inverse peak voltage	3.14	1.57	2.09	2.09	2.09	1.05
3. First three alternating components of rectifier output, rms:						
a. At ripple frequency	0.472	0.472	0.177	0.177	0.040§	0.040
b. Second harmonic of ripple frequency	0.095	0.095	0.040	0.040	0.010	0.010
c. Third harmonic of ripple frequency	0.040	0.040	0.018	0.018	0.004	0.004
B. Ripple frequency (F is supply frequency)	2F	2F	3F	3F	6F	6F
C. Ratio of tube currents to direct current:						
1. Peak anode current	1.00	1.00	1.00	1.00	0.50	1.00
2. Average anode current	0.50	0.50	0.33	0.33	0.167	0.33
D. Transformer utilization factors:						
1. Primary	0.900	0.900	0.827	0.827	0.955	0.955
2. Secondary	0.637	0.900	0.675	0.675	0.675	0.955

See Appendix D for methods of determining the numerical values given in this table.

* This table neglects the drop through the choke and tubes and all reactance drop in the transformers. The term "direct output voltage" as used in this table is, therefore, the equivalent no-drop voltage, E_{n}, referred to in the example, p. 198.

† Voltage on one side of center tap.

‡ Voltage of one coil—two coils required per leg.

§ The fundamental component of voltage across the balance coil has an rms amplitude of 0.354 at a frequency of 3F.

approximately sinusoidal, and higher frequency components must be considered in addition to the fundamental component I_1. These may produce a considerable increase in I_p.

Determination of the peak-current demanded of a tube is of considerable importance, since all commercial tubes are limited in their peak current-carrying capacity. A tube may be able to pass the average current demanded of it without overheating and yet not be able to supply the peak current required. A rectifier designed for operation with such a tube would be unable to deliver the expected output.

Maximum inverse peak voltages are also given in Table 7-1. When a tube is passing current, the voltage impressed between the plate and filament is that due to the drop in the tube itself; but when the tube is nonconducting, the voltage across the tube may be over three times the direct output voltage. Tubes designed to withstand the inverse peak voltage must be selected, or a circuit with a lower inverse peak voltage used.

Complete Rectifier Design. The following example illustrates the application of the foregoing principles in the design of a high-power, high-voltage rectifier. High-vacuum tubes and the double-Y circuit are assumed as representing the more complex type of solution. Solution of a circuit using mercury-vapor tubes would differ only in the use of a constant tube drop of about 12 to 15 volts (regardless of the current flowing) and probably in the use of the bridge circuit of Fig. 7-16. The design of small rectifiers, such as used in radio receivers, may be handled in the same manner, but tube manuals and other sources of information are now available that make the design of most such units a simple matter of picking values out of tables or charts.

Example. Suppose that a rectifier is to be designed having the following specifications, using high-vacuum tubes:

Output	20 kw
Full-load terminal volts.	7000
Maximum ripple in output	½ per cent
Maximum impedance to load.	1000 ohms at 30 cycles

For a rectifier of this large capacity using high-vacuum tubes the double-Y circuit would almost certainly be used, and inspection of Table 7-1 will disclose that E_1 is 0.040 times the direct voltage, or 280 volts. (Since Table 7-1 assumes no drop in the chokes and tubes, the results in the first part of this design will be approximations; but exact figures are not necessary

in the selection of condenser, chokes, and tubes.) The ripple voltage E_R is ½ per cent of the direct voltage, or 35 volts. The ripple frequency is given by the table as six times the frequency of the supply, or 360, so that ω is 2260. Substitution of these values in Eq. (7-3) give a minimum LC product of 1.57×10^{-6}, where L is in henrys and C is in farads. A preferable procedure is to express C in microfarads, in which case $LC = 1.57$.

The minimum value of the capacitance is determined from the last item in the specifications and Eq. (7-4). A condenser having a reactance of 1000 ohms at 30 cycles is found to be of approximately 5 μf capacitance.

The minimum value of L is next solved from Eq. (7-5), assuming I_{min} to be the quarter-load current of 0.715 amp. The minimum L is found to be 0.245 henry.

The product of minimum L and minimum C is 1.23, less than the minimum LC product obtained from Eq. (7-3). It is, therefore, necessary to determine the most economical combination of L and C that will produce a product of 1.57 and still exceed the individual minimum values determined for the inductance and capacitance. The most economical combination usually requires the use of the minimum capacitance and the corresponding inductance as determined by Eq. (7-3). In the problem at hand, therefore, C will be 5 μf and L about 0.32 henry.

The cost of the inductance will depend quite largely upon the amount of resistance permissible. Evidently careful design would require that a balance be obtained between the interest on the investment in the inductance and the cost of the energy wasted in heat losses over a given period of time. Actually the selection is commonly made on the basis of experience and general design data. For the problem at hand assume the 60-cycle power factor of this choke to be 22½ per cent; the resistance will then be 27 ohms.

The same design considerations apply to the balance coil as to the main choke; *i.e.*, the inductance of this unit should be such that the instantaneous current flowing through its windings never drops to zero. The direct current flowing in each half of this unit is obviously half the load current so that the crest value of the alternating current through this choke should not exceed 0.358 amp. Table 7-1 gives the voltage across the coil as 0.351 times the direct voltage, or $0.354 \times 7000 = 2480$ volts, at a frequency of three times that of the supply ($\omega = 1130$). The inductance is, therefore, equal to $2480/(0.707 \times 0.358 \times 1130) = 8.88$ henrys, and, at a 2 per cent 60-cycle power factor, the resistance is 72.8 ohms. Although this inductance is over twenty times that of the filter choke, the cost per henry is very much less since the d-c magnetization in the two halves balances, thereby preventing any tendency to saturate. An air gap must be used in the filter choke to avoid saturation, and many more turns of wire are required to obtain a given inductance than in the balance coil. Consequently the balance coil may cost less than the filter choke in spite of its larger inductance. The lower power factor assumed is also due to this difference in design since the number of turns of wire per henry is smaller than in the choke.

Selection of the tube to be used may now be made. The factors involved in this selection are average current, peak current, and inverse peak voltage.

According to Table 7-1 the inverse peak voltage is 2.09 times the direct voltage, or 14,630 volts; and the average current is 0.167 times the direct load current, or 0.48 amp. The peak current is given as 0.5 times the direct current, or 1.43 amp, but is based on the assumption of an infinite inductance, or $I_1 = 0$. The true peak current is therefore obtained from Eq. (7.7), giving a corrected value of

$$0.5(2.86 + 0.55) = 1.71 \text{ amp}$$

With the foregoing data at hand the most economical tube may be selected.

The only item of design remaining is the power transformer. As this must be capable of producing the full 7000 volts at the terminals of the rectifier, it is necessary at this point to correct for the drop in chokes and tubes. By adding the full load drop through these parts of the circuit to the 7000 volts terminal voltage, an equivalent no-drop voltage E_n is secured. This latter is the voltage on which Table 7-1 is actually based, but obviously it could not be determined until chokes and tubes were selected.

The d-c drop in the choke at the full load is

$$2.86 \times 27 = 77 \text{ volts}$$

while the drop in the balance coil is

$$2.86/2 \times 72.8/2 = 52 \text{ volts}$$

Half the current and half the balance coil resistance are used for reasons obvious from inspection of the circuit. The tube drop is determined from a curve of plate current vs. plate voltage for the type of tube selected. Suppose that it is found to be 750 volts. Neglecting the drop in the transformer (this drop is usually smaller than any of the others under consideration), the equivalent no-drop voltage E_n will be $7000 + 129 + 750 = 7879$ volts. Referring again to the table, the transformer secondary voltage is found to be 0.855 times the direct voltage, or $0.855 \times 7879 = 6740$ volts. The power input to the tubes and filter is E_nI_0 or $7879 \times 2.86 = 22{,}500$ watts. (Since the output was 20,000 watts, the difference of 2500 must represent the loss in tubes and chokes.) The utilization factors of the transformers are given in the table as 0.955 for the primaries and 0.675 for the secondaries, so that the ratings of the windings will be 7.9 and 11.1 kva, respectively, per transformer.

The actual inverse peak tube voltage may now be found as 2.09 times E_n or 16,400 (instead of the approximate value of 14,630 previously found).

Reasonably accurate characteristic curves may be readily obtained for all loads above about one-quarter full load. Except for a very slight transformer drop, E_n remains constant above about one quarter load (the filter inductance becomes ineffective below this point which results in a rise in E_n), so that the terminal voltage is found for any given load by subtracting the choke and tube drops from E_n. The choke drops are obtained by a simple application of Ohm's law, but the tube drop must be read from the characteristic curves of the tube. The efficiency may be determined by dividing the output E_oI_o by the input E_nI_o + filament loss + transformer losses. The filament loss is equal to the number of tubes times E_fI_f for one

tube, since all tubes are drawing filament current continuously. The transformer losses can generally be estimated with sufficient accuracy from tables in handbooks or from specification sheets supplied by manufacturers.

Pi-section Filters. The design of a pi-section filter is more difficult than that of an L-section filter owing to the action of the input condenser. The voltage across this condenser has the general shape of the heavy line in Fig. 7-36 when used with a single-phase, half-wave rectifier. At time t_1 the voltage of the transformer secondary (neglecting tube drop) becomes equal to the voltage of the condenser, and current flows through the rectifier tube, charging the condenser as well as supplying the load. At time t_2 the

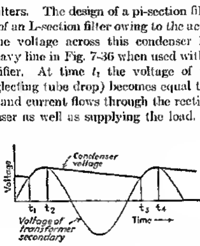


FIG. 7-36.—[illegible] curve for the input condenser C_0 of a pi-section filt[illegible] inductance, no tube drop, and a single-phase, half-wave [illegible]

supply voltage has begun to drop and again becomes less than the voltage of the condenser. The rectifier tube ceases to conduct, and the condenser supplies the load current from t_2 to t_3 when the supply voltage once more equals that of the condenser and the cycle is repeated. The discharge curve will be a straight line if it is assumed that the inductance is large enough to maintain the current constant (a reasonable assumption considering the very short interval of time between condenser chargings), and its slope will obviously be a function of the load current.

The ordinates of this curve may be determined with a fair degree of accuracy by assuming a condenser charge-discharge curve as in Fig. 7-37. This will give a condenser capacitance a little larger than necessary, which is erring on the desirable side. The maximum voltage across the condenser will be equal to the crest value of the alternating voltage of the secondary winding, neglecting tube drop and transformer reactance. The drop in voltage during the discharge period multiplied by the capacitance of the condenser

will be equal to the quantity of electricity required by the load during the same interval of time.

$$(\Delta E_c)C_0 = I_0 T \tag{7-8}$$

where ΔE_c = drop in voltage across condenser

C_0 = condenser capacitance

I_0 = direct load current

T = interval of time between condenser chargings

The time T would be that of one cycle for a half-wave, single-phase rectifier, one-half cycle for a double-wave, single-phase rectifier; one-third cycle for a half-wave, three-phase rectifier; etc.

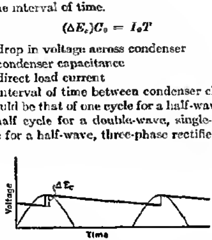


FIG. 7-37.—Simplified charge-discharge curve which closely approximates the true curve of Fig. 7-36

After the condenser-voltage wave of Fig. 7-37 has been evaluated from Eq. (7-8), it may be analyzed by standard Fourier analysis methods,[1] and the various components applied to the remaining portion of the circuit which is merely an L-section filter. Normally only the fundamental component (or ripple frequency) need be considered. As a matter of fact, design of a pi-section filter is not often required, as this type is commonly used only in low-power rectifiers where careful design is not necessary owing to the low cost of the component parts

A resistor is sometimes inserted in series with the rectifier output but ahead of the filter when a condenser-input filter is used. This protects the tubes against excessive current when the rectifier is started If the a-c input to the rectifier has been deenergized for a sufficient length of time for the condenser to become discharged through the resistance of the load circuit and the alternating current is reapplied, the condenser will act as a virtual short circuit until it begins to charge. The series resistor should be of such size as to limit the current into the condenser to the safe rating of the

[1] Appendixes B and C.

tubes, even though its presence will somewhat reduce the efficiency of the rectifier.[1]

In low-power rectifiers the cathodes of the tubes are often heated from an extra winding on the high-voltage supply transformer and will therefore be cold at the time the alternating current is applied. The time lag in heating of the cold cathodes is sufficient to prevent excessive currents from flowing until the condenser has become charged. Where this is not so, either a resistor should be used or the filter should be of the inductance-input type.

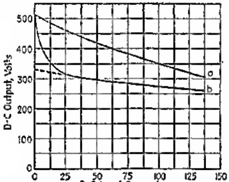


FIG. 7-38.—Output voltage curves of a low-voltage, single-phase, full-wave rectifier with (a) pi-section filter and (b) L-section filter.

Operating Characteristics. The regulation curve of a full-wave, single-phase, low-power rectifier is shown in Fig. 7-38 with a pi-section filter and an L-section filter. It should be noted that the regulation is very much better with L-section filter except at very low loads but that the output voltage is considerably less. When a condenser input is used, reducing the load current will cause the slope of the discharge curve, from t_2 to t_3 in Fig. 7-36, to decrease, thereby shortening the interval t_3t_2 of current flow through the tube whence the average condenser voltage will approach the crest value of the impressed alternating voltage as a limit. The output voltage of the pi-section filter at no load is seen to be 515 volts which is essentially the crest value of the impressed emf. (The secondary voltage was 370 volts rms.)

[1] A. P. Kauzmann, Determination of Current and Dissipation Values for High-vacuum Rectifier Tubes, *RCA Rev.*, 8, p. 82, March, 1947.

In the L-section filter (Fig. 7-22), with the condenser C_0 of Fig. 7-21 eliminated, the inductance causes the tube current to be nearly constant so that the output voltage should approach the average value of the rectified emf (333 volts) as the load current approaches zero. It may be seen that curve (b) of Fig. 7-38 does tend to approach this point but finally curves abruptly upward toward the same voltage at no load as was obtained for the pi-section filter. The inductance becomes noneffective at very light load currents, and the condenser C then charges up toward the crest value of the applied emf exactly as did C_0 in the pi-section filter.

A serious objection to this abrupt rise in voltage at light loads is that the condenser C must be designed for the maximum voltage[1] of 515, whereas during normal operation of the rectifier the voltage never exceeds about 325. A bleeder resistance may be connected across the output terminals of inductance-input filters of such size as to prevent the current through the filter from becoming so small that the inductance loses its effectiveness, thus permitting the use of a lower voltage, cheaper condenser than would otherwise be possible. In many cases the presence of a voltmeter across the output is sufficient to prevent an excessive voltage rise. The use of polyphase circuits also considerably reduces this trouble, as the ripple voltages applied to the filter are much less than for the single-phase case (see Fig. 7-31).

Terminal-voltage and efficiency curves of a 15-kw, 5000-volt rectifier using a three-phase bridge circuit with L-section filter and mercury-vapor tubes are shown in Fig. 7-39. The voltage regulation is much better than for the small rectifier of Fig. 7-38, since the tube and choke drop are a much smaller percentage of the terminal voltage. The efficiency is seen to be quite high, 94 per cent at the maximum. Had high-vacuum tubes been used, the efficiency would have been of the order of 75 to 80 per cent, and the voltage drop, between one-fourth and full load, would have been perhaps 12 to 15 per cent of the rated voltage, even when using the double-Y circuit.

Rectifiers Using R-F Power.[2] High-voltage rectifiers supplied by 60-cycle power require relatively costly power transformers.

[1] This voltage is determined, when designing a rectifier, by multiplying the rms secondary voltage, as found from Table 7-1, by $\sqrt{2}$.

[2] For further information on this type of rectifier, see O. H. Schade, Radio-frequency-operated High-voltage Supplies for Cathode-ray Tubes,

When the power output required is low, as in the rectifiers used to supply cathode-ray tubes, it may be more economical to obtain direct voltages of the order of 10,000 or more by rectifying a r-f source obtained from a vacuum-tube oscillator (see Chap. 11). A r-f source of alternating voltage permits the use of a tuned, air-core transformer instead of the usual iron-core type, at a considerable saving in cost. It also reduces the cost of the filter as Eqs. (7-3), (7-4), and (7-5) show that the size of the inductance and condenser required for filtering varies inversely with the frequency. Since the saving in transformer and filter costs is at least partly offset by the added cost of the vacuum-tube oscillator, power supplies of this type are economical for low power only, where the added cost of the oscillator is less than the savings in the rectifier.

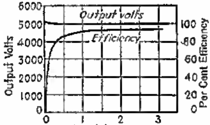


FIG. 7-39.—Terminal voltage and efficiency curves of a 15-kw., 5000-volt rectifier using a three-phase bridge circuit and mercury-vapor tubes.

It is even possible to combine the action of the rectifier and r-f oscillator in a single tube. As will be demonstrated in Chap. 11, vacuum-tube oscillators generate their own direct voltage for biasing the grid, and it is possible, with proper tube and circuit design, to generate a direct voltage in the grid circuit which is much greater than the direct voltage supplied to the plate.[1] The direct plate voltage is supplied by a conventional rectifier, the voltage of which is low enough to offer no problems in power transformer design.

Proc. IRE, 31, p. 158, April, 1943; Robert S. Mautner and O. H. Schade, Television High Voltage R-f Supplies, *RCA Rev.*, 8, p. 43, March, 1947.

[1] Robert L. Freeman and R. C. Hergenrother, High-voltage Rectified Power Supply Using Fractional-mu Radio-frequency Oscillator, *Proc. IRE*, 34, p. 145W, March, 1946.

Voltage-multiplying Circuits.[1] Figure 7-40 shows a full-wave rectifier circuit in which the direct no-load voltage will be twice the crest value of the emf across the secondary. It is very similar to the bridge circuit (Fig. 7-7), except that two of the tubes have been replaced by condensers.

In order to explain the operation of this circuit it will be assumed that it has been in operation sufficiently long for all transients to have disappeared. During the half cycle that the left-hand terminal of the secondary winding is positive, current will flow through tube 1, through the load circuit, and back to condenser C_2. As will be seen shortly, condenser C_2 will, at the beginning of the half cycle, be charged to a potential somewhat less than the crest value

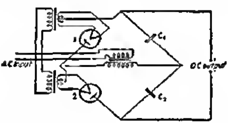


FIG. 7-40.—Voltage-doubling circuit. The direct output voltage approaches twice the crest alternating voltage.

of the secondary voltage. During the half cycle under discussion it will discharge, adding its potential to that of the secondary and so increasing the potential delivered to the external circuit. At the same time condenser C_1 will be charged through a circuit consisting of the transformer secondary and tube 1. In the next half cycle, the right-hand terminal of the secondary will be positive, and C_1 will discharge, causing a flow of current through C_1, the load circuit, tube 2, and back through the secondary winding. At the

[1] For more detailed information on the performance of voltage multiplying circuits the reader is referred to the following: D. L. Waidelich, The Full-wave Voltage-doubling Rectifier Circuit, *Proc. IRE*, 29, p. 554, October, 1941; D. L. Waidelich and C. H. Gleason, The Half-wave Voltage-doubling Rectifier Circuit, *Proc. IRE*, 30, p. 535, December, 1942; D. L. Waidelich and C. L. Shackleford, Characteristics of Voltage multiplying Rectifiers, *Proc. IRE*, 32, p. 470, August, 1944; D. L. Waidelich and H. A. T. Taskin, Analyses of the Voltage-tripling and -quadrupling Rectifier Circuits, *Proc. IRE*, 33, p. 449, July, 1945.

same time, condenser C_2 will be charged, and the cycle will be repeated.

Figure 7-41 shows the circuit of a cascade doubler (also known as a half-wave doubler) which permits grounding of one terminal of both a-c input and d-c output circuits since, unlike the circuit of Fig. 7-40, one terminal of both circuits is common. When terminal B of the a-c input is positive, current flows through tube

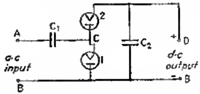


FIG. 7-41. Circuit of a cascade (or half-wave) doubler.

1 charging condenser C_1 to a potential which approaches the crest value of the a-c supply. On the next half cycle, when terminal A is positive, the voltage of condenser C_1 adds to that of the supply causing current to flow through tube 2 and charging condenser C_2 to a potential equal to the sum of the supply voltage and that of condenser C_1, a total approximately equal to twice the crest value

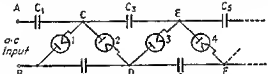


FIG. 7-42. Illustrating the method of expanding the circuit of Fig. 7-41 to any desired degree of multiplication.

of the supply voltage. Since condenser C_2 must supply the load during a large part of the cycle between successive charges, the *average* voltage across its terminals will be somewhat less than twice the crest value of the supply voltage, dropping with the load in a manner similar to the operation of the condenser-input filter as described on page 199.

Figure 7-42 shows how the circuit of Fig. 7-41 may be expanded to produce any desired degree of multiplication. If terminal B is used as the negative terminal of the d-c output, the potential at D will approach twice the crest value of the applied alternating volt-

age, the circuit $ABCD$ being that of Fig. 7-41 where the same notation was used. Similarly the potential at F will approach four times the crest value of the a-c voltage. If terminal A is used as the negative d-c terminal, the voltage at C will approach the crest value of the applied alternating voltage, and the voltage at E will approach three times the crest value. This process may evidently be continued to higher voltages by adding more tubes and condensers.

Voltage-multiplying circuits are used where either the direct voltage required is so high that it is not economical to build a transformer of sufficient voltage for one of the conventional half-wave or full-wave circuits, or where the direct voltage is to be obtained directly from the a-c supply without the use of a transformer, as in certain types of radio receivers. In the latter application the cathode heaters of the tubes are frequently designed to operate at 25 or 35 volts, whence three or four heaters are connected in series with each other and with a suitable resistor and bridged across the 110-volt a-c supply line.

The use of voltage-multiplying circuits should be avoided unless the output current is quite small, because the regulation is poor. As may be clearly seen from Fig. 7-42, the higher voltages are obtained by essentially charging condensers in parallel and then discharging them in series to supply the output. Unless unreasonably large condensers are used, the voltage will drop rapidly with increasing load for the reasons given in explaining curve a of Fig. 7-38 for the condenser-input filter. Much larger condensers must be used than in the condenser-input filter since several are connected in series thus reducing the effective capacitance. As between the circuits of Fig. 7-40 and 7-41, the former has slightly better regulation and lower ripple while the latter has a common input and output terminal for grounding, an important consideration when no power transformer is used.

Special Circuits for Use with Ignitrons. In general, ignitrons may be used in the same circuits as hot-cathode, mercury-vapor tubes. Nevertheless certain auxiliary equipment is necessary to fire the igniter.

The method by which this firing is accomplished is illustrated by the simple, single-phase, half-wave circuit of Fig. 7-43. A small, hot-cathode, mercury-vapor rectifier B is used to provide satisfactory operation of the igniter I. In Fig. 7-41 the applied

potential is shown by the sine wave—partly dotted—of curve (*a*). As this potential starts rising in a positive direction after the time indicated at point 1, it finally reaches a potential at point 2 where

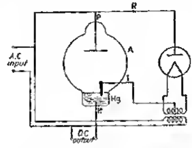


FIG. 7-43.—Circuit of an ignitron tube and controlling rectifier.

tube *B* breaks down and passes current as shown in curve (*c*). This current rises with the applied potential until sufficient to cause an arc between the igniter electrode and the mercury pool, at point 3. The anode potential (heavy line of curve *a*) is now sufficient to draw the arc over to the anode, causing anode current to flow as in curve (*b*) and the potential drop within the ignitron to fall immediately to about 15 volts. Since the drop within tube *B* is of this same order of magnitude and the drop at the igniter electrode is about an additional 10 volts, the potential between points *P* and *K* is insufficient to maintain current flow through tube *B*, and the current through the igniter circuit falls to zero, remaining zero as long as the ignitron continues to pass current.

FIG. 7-44.—Currents and voltages in the circuit of Fig. 7-43. (*a*) Voltage between anode and cathode of the ignitron. (*b*) Current flowing through the ignitron. (*c*) Current flowing through the igniter circuit.

As the applied potential again passes through zero at the end of the positive half cycle, the current flow through the ignitron ceases. The presence of the small rectifier *B* prevents a reverse

current flow through the igniter circuit, and consequently there is somewhat less danger of flashback during the negative half cycle than in the mercury-vapor, hot-cathode type of tube and very much less than in the mercury-arc tube where ionized mercury vapor is present at all times throughout the cycle. The tube, therefore, remains entirely inactive until the applied potential again reaches the positive half of its cycle, when the process described in the preceding paragraph is repeated.

The application of this type of rectifier to any of the circuits previously described in this chapter should be obvious. The only change required is the addition of the small rectifier for controlling the igniter circuit of each ignitron.

The current required for firing is comparatively large, being from 10 to 50 amp, but the total energy required is quite small because of the short duration of the firing pulse, generally of the order of microseconds. For this same reason the capacity of the controlling rectifier need not be so large as if it were required to carry a sustained current of such magnitude. Mercury-vapor, hot-cathode rectifiers are capable of safely passing currents of two to five times their continuous ratings when the duration of the pulse is much less than 1 sec. Nevertheless the demand on these tubes is severe, and special tubes have been designed for this service which have a ratio of peak-to-average current as high as 10 or 12.[1]

Special Igniter Circuits to Ensure Firing. One of the major problems in the use of ignitrons is to make certain that the anode will pick up the discharge as soon as the igniter strikes. If it does not, the current through the igniter will continue to flow until the anode potential does finally become sufficiently high to draw the arc. This imposes such a heavy duty on the control rectifier and igniter electrode as to shorten their life greatly. It is therefore necessary to ensure absolutely that the anode will pick up the current flow as soon as the arc is struck. Where difficulty is experienced, insertion of a resistance at point R in Fig. 7-43 will ensure satisfactory results. The anode potential at the moment of ignition is equal to the total drop through the igniter circuit. The resistance R evidently increases this drop and so delays ignition until the anode potential has reached a somewhat higher value, say point 4 in Fig. 7-44. Since ignition takes place at a later time

[1] D. D. Knowles, E. F. Lowry, and R. K. Gessford, A New Welder Tube, *Electronics*, 9, p. 27, November, 1936

in the cycle, the average anode current will be less; therefore, it is desirable to keep the resistance R as low as possible if the full capacity of the tube is to be realized.

Another method that sometimes proves satisfactory is the insertion of an inductance at the point R (Fig. 7-43). The inductance will maintain the potential across the igniter circuit for a brief time after ignition takes place, regardless of the potential difference across the ignitron itself, and so increase the probability of the anode picking up the current flow.

Other Types of Firing Circuits. Elimination of the tube B from the ignition circuit would seem to be a desirable development, since its life is limited under the high peak currents that it must pass in this service. Many attempts have been made to devise a circuit without vacuum tubes that will ensure ignition and yet interrupt the current to the igniter after ignition and prevent its flow during the negative half cycle, the most promising of which involve the use of saturated iron cores. It is well known that the current flow through a saturable reactor, with sinusoidal impressed emf, is highly peaked as in Fig. 7-45a. This is due to operating over the range x of the hysteresis curve (c). Such a peak of current may be sufficient to ignite the arc in the ignitron, but unfortunately the same peak occurs during both positive and negative half cycles and would probably cause arc back. If a suitable direct current is also passed through the circuit, however, the negative peak may be eliminated, as in (b) of Fig. 7-45, by confining operation to the region y of the hysteresis loop (c).

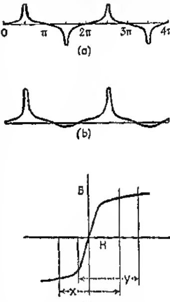


FIG. 7-45.—Principle of operation of an ignitron firing circuit using a saturated iron core.

In practice, the saturable reactor is connected in series with the

igniter of the tube, and a direct current is supplied by a small rectifier, usually of the copper oxide type.[1] Care must be exercised to supply the proper phase of alternating current to the igniter circuit to ensure ignition at the desired point of the cycle. This can be done by proper transformer connections from a three-phase

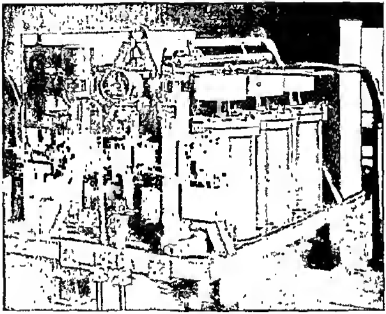


FIG. 7-46. Ignitron rectifier for supplying power to the underground transportation system of a large mine in the western part of the United States. (*Westinghouse Electric Corp.*)

circuit or by means of one of the phase-shifting circuits described in Chap. 8.[2]

Figure 7-46 shows a full-wave, three-phase ignitron rectifier used

[1] For a discussion of this type of rectifier, see J. Slepian, Thin Film Rectifiers, *Trans. Am. Electrochem. Soc.*, 54, p. 201, 1928; also L. O. Grondahl and P. H. Geiger, A New Electronic Rectifier, *Trans. AIEE*, 46, p. 357, 1927.

[2] For detailed circuits see Hans Klemperer, A New Ignitron Firing Circuit, *Electronics*, 12, p. 12, December, 1939.

to supply the d-c power for the underground transportation system of a large Western mine. Figure 7-47 shows a similar rectifier which is fully portable and may be moved around inside the mine as needed to keep the d-c transmission distances low. The individual ignitron cells are readily distinguishable in both figures by the cooling fins on the anode terminals at the top of the metal envelopes.

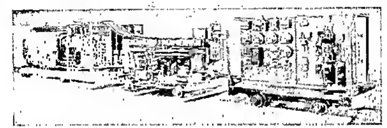


FIG. 7-47.—Portable underground 300-kw. automatic ignitron rectifier; 4000 volts, 3 ϕ, to 275 volts, d-c. (*Westinghouse Electric Corp.*)

Special Circuits for Mercury-arc Rectifiers. In general, mercury-arc rectifiers are of the polyphase type and are used with the double-Y circuit of Fig. 7-13 or the full-wave circuit of Fig. 7-11, the bridge circuit of Fig. 7-16 being unsuitable owing to the use of a single cathode in a polyphase mercury-arc rectifier. Figure 7-48 shows the connections for a mercury-arc tube in a double-Y circuit. In addition auxiliary circuits must be supplied to start the arc, and a keep-alive circuit or other means must be provided to prevent extinguishing of the arc during momentary cessations of current. The operation of these auxiliary circuits may best be illustrated as applied to a simple single-phase circuit using the now obsolete glass-envelope mercury-arc tube of Fig. 4-11.

The circuit for this tube is shown in Fig. 7-49. It is essentially the same as that of Fig. 7-6 except for the auxiliary circuits required for starting and maintaining the arc.[1] The arc is started by tipping the bulb until the mercury from the main pool runs over into the starting pool S, closing the circuit through the current-limiting resistance R and one-half the secondary winding of the auxiliary transformer T_2. The bulb is then returned to its normal position, and the arc is formed as the pools separate. The

[1] See p. 108 for methods of starting.

arc quickly jumps to the keep-alive anodes K_1, K_2, and current then flows through one-half the auxiliary transformer T_2 and anode K_1 to the cathode C during one half cycle of the impressed alternating voltage and through the other half of the transformer and the other anode K_2 during the second half cycle. This entire circuit is termed the *keep-alive*, and its purpose is to maintain sufficient ionization of the gas within the tube to permit stable operation of the main circuit through anodes P_1 and P_2 regardless

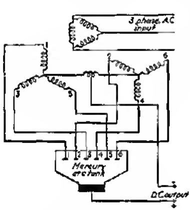


FIG. 7-48.—Circuit of a six-anode, mercury-arc rectifier, using the double-Y circuit.

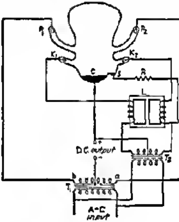


FIG. 7-49.—Circuit of a two-anode, mercury-arc rectifier.

of the load. The voltage is low, and the energy consumed in this circuit is comparatively small. Nevertheless it does constitute an appreciable loss in low-voltage, low-power rectifiers.

The flow of current through the tube is unidirectional, though pulsating, since the anodes attract electrons but do not emit them. The temperature of the anodes is kept sufficiently low to prevent the production of electron emission at their surfaces by thermionic action; and since they are located at the ends of long arms, there is little tendency for the slower moving positive ions to reach them and so produce a reverse current.

The voltages of the secondary of transformer T_1 are shown in Fig. 7-50, curve (*a*); and the current that flows through anodes

P_1 and P_2, neglecting transformer leakage reactance and assuming a resistive load, is shown in curve (b), the tube drop being assumed constant throughout the conducting period. It will be seen that there is a period of time t_1t_2 during which no current flows. If it is assumed that no keep-alive circuit is being used, *i.e.*, K_1 and K_2 are disconnected, the arc would be extinguished at time t_1 when the current went to zero, and the tube would stop functioning. This may be remedied by inserting inductance in the circuit so that the current in one anode will be maintained by the back emf of the inductance until the other anode starts to pass current. The current flowing through the cathode lead will then have the appearance of curve (c). This inductance may be supplied by leakage reactance in the transformer or by the insertion of an additional inductance in series with the output circuit. Small-sized rectifiers are sometimes built without keep-alive anodes and operated in this manner.

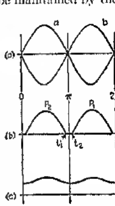


FIG. 7-50.—Voltages and currents in the circuit of Fig. 7-49. (a) Secondary voltages of transformer T_1; (b) anode current without series inductance; (c) anode current with series inductance.

The insertion of inductance into the load circuit is often undesirable. For this reason all large-size rectifiers use a keep-alive circuit which must include sufficient inductance to maintain the arc. (In Fig. 7-49 the latter inductance is provided by the unit L.)

Polyphase mercury-arc rectifiers may be operated without a keep-alive circuit and without inductance in series, since the voltage on one anode does not drop to zero until after another anode has picked up the load. However, keep-alive circuits are normally used to guard against extinguishing of the arc by momentary removal of the load, although only sufficient auxiliary equipment to operate one pair of keep-alive electrodes and the starting circuit is needed.

Operating Characteristics. The regulation of these rectifiers is determined primarily by the transformers used, since the tube drop is virtually independent of the amount of load being carried. Their efficiency is also quite high, ranging as high as 90 to 97

per cent depending on the voltage and power rating, the higher efficiencies being obtained at high voltage and power outputs. The lower efficiency at low voltages is due largely to the fixed tube drop of 15 to 40 volts. At low output voltages this constitutes a rather high percentage of the total drop, whereas at high voltages it may become practically negligible. At low-power outputs the keep-alive circuit drops the efficiency, since it draws a continuous load of 100 watts or more regardless of the total power delivered by the tube. This is, however, considerably less than the power required to heat the cathodes of a bank of large-capacity mercury-vapor tubes.

Controlled Rectifiers. It is often desirable to vary or arbitrarily adjust the voltage of a rectifier between certain limits. In experimental work, for example, a rectifier capable of delivering any voltage from zero to its maximum is often very useful. One means of so doing is to provide separate transformers for the plate and filament supplies with a variable voltage a-c source feeding the plate transformer. Another method is to use grid-controlled rectifier tubes in which the amount of current passing, and therefore the output voltage, is controlled by varying the potential applied to the grid. Still another is to insert a grid-controlled tube in series with the d-c output to control the amount of current flowing by varying the effective series resistance. The latter two methods are discussed in the ensuing chapter (see pages 225 and 237).

Problems

7-1. The output of a full-wave, single-phase rectifier similar to Fig. 7-3 is filtered by a single L-section filter. The curve of tube drop for each tube is given by curve 1, Fig. 3-3. The transformer secondary voltage is 450 volts rms each side of the center tap. The filter inductance has a resistance of 150 ohms. Calculate and plot a curve of direct output voltage vs. direct load current from 0 to 150 ma. (Assume infinite L and zero transformer reactance.)

7-2. Repeat Prob. 7-1 for the bridge circuit of Fig. 7-7. The transformer secondary voltage is 450 volts rms.

7-3. A half-wave, single-phase rectifier (Fig. 7-1) is to be designed and will be used for supplying grid bias to an amplifier. The filter consists of a condenser only, and the load is a resistance of 250,000 ohms bridged across the terminals of the rectifier. The desired bias voltage of 300 volts is set up across this resistor. If the maximum change in voltage across the condenser is not to exceed ½ per cent of the output voltage, what should be the capacitance of the filter condenser? Assume amplifier grid current is zero

7-4. The output voltage of a given rectifier is to be 10,000 volts at full load, and an L-section filter is to be used. What voltage is impressed across the filter condenser at no load when using the following circuits: (a) Fig. 7-1, (b) Fig. 7-3, (c) Fig. 7-7, (d) Fig. 7-8, (e) Fig. 7-10, (f) Fig. 7-11, (g) Fig. 7-13, (h) Fig. 7-16? (Neglect drop in tubes, choke, and transformers.)

7-5. Compute all the values called for in Table 7-1 for the circuit of Fig. 7-11. (Follow the method of Appendix D.)

7-6. (a) Solve for the second and third harmonics of the ripple current flowing through the inductance of the example on page 197. (b) What voltage does each harmonic of part (a) set up across the output terminals?

7-7. A full-wave, single-phase rectifier has a transformer secondary voltage of 375 rms each side of the center tap. It is to deliver a current which will vary from 10 ma minimum to 50 ma maximum. Neglecting tube and transformer drop (a) what is the minimum inductance required for an L-section filter? (b) What is the peak current per tube at maximum load? (c) What is the inverse peak voltage per tube?

7-8. A single-phase, full-wave rectifier has a secondary voltage of 375 rms each side of the center tap with a full load rating of 20 ma. The output is filtered by a two-section, inductance-input filter like that of Fig. 7-30. If C_1 and C_2 in this figure are each 4 μf and L_1 and L_2 are each 8 henrys, (a) what is the per cent ripple in the output? (b) What is the per cent ripple if the inductances and capacitances are lumped into a single-section filter such as that of Fig. 7-22, where L will then be 16 henrys and C, 8 μf? (c) Which circuit provides the more effective filtering?

7-9. The rectifier of Prob. 7-8 is to be operated into the pi-section filter of Fig. 7-21, where C_0 and C are 4 μf each and L is 16 henrys. (a) Compute ΔE_s of Fig. 7-37 at full load. (b) Determine the fundamental component of the resulting voltage across C_0, following a procedure similar to that of Appendix C. (c) Compute the per cent ripple due to the fundamental component of the ripple voltage across C. (d) Compare these results with those of Prob. 7-8, where the same total inductance and capacitance were used. (Note this is not an entirely fair comparison since there are rather large higher order harmonics present here.)

Design Problems[1]

7-10. Design a rectifier from the following specifications, using a three-phase, double-Y circuit, L-section filter, and high-vacuum tubes. Assume the tube-drop curve to be given by $I_b = 12.5 \times 10^{-6} \times (E_b)^{3/2}$ amp, where E_b is in volts.

Specifications:

Output in kilowatts	10
Full-load terminal voltage	5000
Maximum ripple in output	½ per cent
Maximum impedance to load in ohms	500 at 40 cycles
Minimum load current	¼ of full load current

[1] The solution of the following problems is somewhat long and assignments should be made accordingly.

Use a filter choke with a power factor of 20 per cent and a balance coil with a power factor of 2 per cent, both taken at 60 cycles. In the design specify [illegible]

the inverse peak voltage, peak current, and average current of the tubes; and compute and plot curves of efficiency and voltage output vs. direct load current. Filament loss, 600 watts per tube. (Assume power and filament transformer efficiencies of 97 per cent. The efficiency will actually vary with load, but little error will result by assuming it constant, except at very light loads.)

7-11. Design a rectifier from the following specifications, using a three-phase bridge circuit, L-section filter, and hot-cathode, mercury-vapor tubes. Assume the tube drop to be 12 volts, independent of the amount of current flowing.

Specifications

Output in kilowatts	10
Full-load terminal voltage	7500
Maximum ripple in output	$\frac{1}{4}$ per cent
Maximum impedance to load in ohms	600 at 30 cycles
Minimum load current	$\frac{1}{4}$ of full load current

Use a filter choke with a power factor of 20 per cent at 60 cycles. Filament loss, 50 watts per tube. In the design specify the same quantities and plot the same curves as called for in Prob. 7-10 (except, of course, for the balance coil). (Assume the same transformer efficiencies as in Prob. 7-10.)

CHAPTER 8

THE VACUUM TUBE AS A CONTROL DEVICE

Insertion of a grid into the vacuum tube gave it the properties of a control device. Thus it is possible to apply a voltage to the grid of a tube and control large amounts of power in the plate circuit with but little power being consumed from the controlling source by the grid circuit.

Grid control may be classified as (1) uniform control and (2) "on-off" control. With uniform control, the current flowing through the plate circuit is at all times under absolute control of the grid, small variations in grid voltage being accompanied by small variations in plate current. Generally speaking, only high-vacuum tubes are capable of performing this type of control. Gas-filled tubes are normally capable of providing only an on-off control; *i.e.*, they will permit either the full flow of current through the plate circuit or none, but cannot graduate the flow between these limits. Actually this is true only for instantaneous values of current and voltage, and gas-filled tubes *can* be used to vary the *average* flow between zero and a maximum, where the average is taken over one cycle of an a-c supply (normally 60 cycles/sec, since gas-filled tubes are commonly used in the control of 60-cycle circuits).

1. HIGH-VACUUM TUBES

The high-vacuum tube may be used to control the current flowing in a given load by means of the circuit of Fig. 8-1. The grid battery voltage is equal to or slightly greater than the cutoff voltage of the tube. Thus no plate current will flow when the potentiometer P is moved to its extreme left position, whereas moving the potentiometer to its extreme right position will increase the load current to a maximum. Other potentiometer positions will allow any desired current flow between these two limits.

The moving arm of the potentiometer of Fig. 8-1 may be operated by means of a delicate control mechanism with which it is

desired to control the current through the load. The potentiometer itself may be wire wound with a sliding contact; it may be a liquid with an immersed contact moving between two fixed plates in the liquid; or it may be some other device capable of varying potential but incapable of handling an appreciable current. The d-c sources for grid and plate voltages may be rectifiers or other suitable sources, instead of batteries as shown.

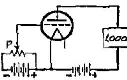


FIG. 8-1.—Circuit of a high-vacuum tube used to control the current flowing through a load.

A-C Operation. The circuit of Fig. 8-1 may be operated entirely from an a-c supply, as in Fig. 8-2. Grid and plate voltages are both supplied from a single transformer tapped at suitable points. The power for heating the cathode may be taken from another winding on the same transformer, not shown. The plate voltage is supplied between points 2 and 3 and is shown by curve 1 of Fig. 8-3*a*. The dotted curve 2 of Fig. 8-3*a* is the cutoff voltage of the tube so that no plate current will flow if the grid

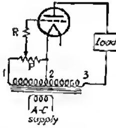


FIG. 8-2.—Circuit of Fig. 8-1 but with an a-c supply.

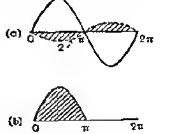


FIG. 8-3.—Curves of voltage and current for the circuit of Fig. 8-2.

voltage is at all times equal to or more negative than this voltage during the positive half cycle of plate supply. The range of voltage variation of the potentiometer is shown by the shaded sections. The plate current will thus flow during the positive half cycles of plate voltage in a magnitude depending upon the position of the potentiometer. Such current flow, for a position

of the potentiometer that provides a grid voltage more positive than cutoff, is indicated by the shaded area of Fig. 8-3*b*.

The resistance R in series with the grid of the tube in Fig. 8-2 prevents the flow of excessive grid current during the positive half cycle of grid voltage. The plate voltage is negative at this time; and as was shown in Chap. 3 (page 69), the grid current is a maximum when the plate voltage is zero (even more so, if negative).

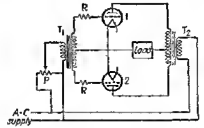


FIG. 8-4. Modification of Fig. 8-2 which provides a more uniform flow of current through the load.

Thus the grid current that would flow without the resistance R might be sufficient to cause damage to the tube or to the potentiometer. The presence of this resistor, even if it is several thousand ohms, will not affect the normal operation of the tube, since the grid is negative at all times during plate-current flow and the grid current will be zero.

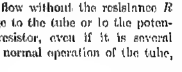


FIG. 8-5. Current flow through the load of Fig. 8-4.

A more uniform flow of current through the load may be secured by the use of two tubes (Fig. 8-4). The plates of the two tubes are energized from opposite ends of a single transformer so that the plate of one will be positive whenever the other is negative, whereas the grids are so energized that each is out of phase with its plate by 180 deg. Thus each tube will perform exactly as did the one in Fig. 8-2 but will pass current during alternate half cycles of the supply. The plate current will then flow in successive half-cycle pulses (Fig. 8-5), the magnitude of the pulses being a function of the position of the potentiometer. The current flowing is therefore a pulsating direct current. A reasonably

steady direct current may be obtained from this circuit by using a condenser in parallel with the load to maintain constant voltage throughout the cycle, or a complete L- or pi-section filter may be inserted, as this unit is actually a full-wave, single-phase rectifier with grid control. (Compare this circuit with that of Fig. 7-3, and note the similarity. The d-c output of the latter figure corresponds to the load of Fig. 8-4.)

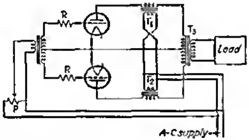


FIG 8 6 —Modification of the circuit of Fig 8-4 to provide a true alternating current through the load

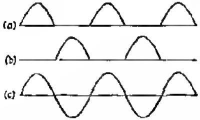


FIG 8-7 —Current flow in the circuit of Fig 8-6, (*a*) and (*b*) show the current flowing in each tube, and (*c*) shows the load current.

It often happens that the load should be supplied with alternating current, rather than with direct current as in Figs 8-2 and 8-4. Figure 8-6 shows a modification of Fig 8-4 by means of which this is accomplished. The plates of the two tubes are again energized 180 deg out of phase but by separate transformers (or by separate windings on the same transformer), and the grids are energized in exactly the same manner as in Fig. 8-4. The load, however, is supplied through a transformer that reverses the polar ity of the emf induced by one plate current with respect to that of the other. The plate current of one tube is shown in curve

(*a*) of Fig. 8-7, and that of the other is shown in (*b*) of the same figure, with curve (*c*) then illustrating the resulting current flowing in the secondary of the transformer T_3, where the half cycles of curve (*b*) have been inverted by the action of the transformer. This secondary current is a true alternating current, the magnitude of which is under complete control of the potentiometer P although, in actual practice, it will not be quite sinusoidal, owing to nonlinearity of the tube characteristic curves.

Operation of Mechanical Relays. It is often desirable to operate a mechanical relay from a set of very delicate contact points or by means of a photoelectric cell or from some other source that will provide a change in voltage with but little power capacity. The vacuum tube is ideal for such purposes, and in many cases either a high-vacuum or a gas-filled tube may be used. If the

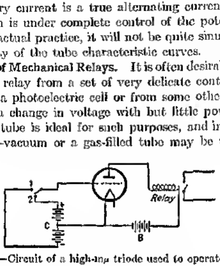


FIG. 8-8.—Circuit of a high-mu triode used to operate a relay.

relay does not require too much current, however, the high-vacuum tube will probably prove cheaper and simpler to operate. For heavy-current relays the gas-filled tube must be used in circuits presented in Part 2 of this chapter.

A simple circuit for operating a relay is that of Fig. 8-8, which is similar to that of Fig. 8-1 except that a switch has replaced the potentiometer, and the relay serves as the load in the plate circuit of the tube. The batteries may be replaced by rectifiers or other d-c sources.

When the switch is thrown to position 1, the grid is made sufficiently negative to prevent the flow of current to the plate, and the relay will not operate. Throwing the switch to position 2 applies a positive voltage to the grid and causes a plate current to flow, closing the relay contacts. The amount of positive voltage applied to a given tube must be sufficient to pass the current required by the relay. The voltage must not, of course, be so high as to endanger the tube or to cause an excessive plate current to flow and so overheat the plate. If the relay requires more cur-

rent than the tube can safely pass, a tube of larger capacity must be used

The amount of power required in the grid circuit of this vacuum-tube relay is not large. The current flowing in the grid, even when made positive, is small. The switch shown may actually consist of very delicate contacts which are incapable of handling any appreciable current.

There is no need to insert a resistor in series with the grid in the circuit of Fig. 8-8 to limit the grid current, as was done in the circuit of Fig. 8-2. The magnitude of the positive voltage that must be applied to the grid of the tube is determined by the magnitude of the plate current required to operate the relay, and this is the voltage which should be supplied by the positive battery. Insertion of a resistor would reduce the grid voltage as well as limit

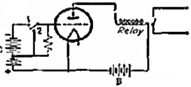


FIG. 8-9.—Circuit of a low-mu triode used to operate a relay.

the current and might prevent operation of the relay. An exception to this statement might be made if the source of supply for the positive potential was unavoidably larger than necessary, whence a resistor might be inserted of such size as to drop the potential to the minimum needed to operate the relay.

The power drawn by the grid of a vacuum-tube relay may be made virtually zero by the use of a low-mu tube. Such tubes require a very high negative voltage in order to secure cutoff; therefore, they will pass a comparatively high plate current with either zero or a low, negative bias. The circuit to be used with such a tube is shown in Fig. 8-9. The C battery is of sufficient voltage to bias the tube to cutoff so that, as before, throwing the switch to position 1 will prevent the flow of plate current. With the switch in position 2, however, a large plate current will flow, even though the grid is still a small amount negative. The current drawn by the grid under these conditions is negligible, since the grid is at all times negative.

An objection to the circuit of Fig. 8-9 is that the grid of the tube is free during the time the switch is moving from position 1 to position 2. As was seen in Chap. 3 (page 54) the free-grid potential is only very slightly negative, and the plate current flowing during the time the switch is not making contact with either point may be excessive. This may be avoided by merely connecting a high resistance between the grid and terminal 2, as shown by the dotted line in Fig. 8-9. It would then be necessary only to open the switch from position 1 to cause the desired plate current to flow. It is true that the use of such a resistor will cause contact 1 to carry a small current, but the resistance can be of the order of hundreds of thousands or even millions of ohms, and thus the current need be but a few microamperes at most.

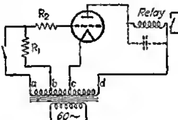


FIG. 8-10.—Modification of Fig. 8-9 using a-c supply.

A-C Operation of Relays. It is frequently more convenient to supply the relay tube with a-c than with d-c power. This may be done by modifying the circuit of Fig. 8-9 to one like Fig. 8-10. Here the secondary of a 60-cycle transformer is tapped to provide plate voltage between points c and d, grid bias sufficient to prevent plate-current flow between points c and a, and grid bias sufficient to pass the desired current through the relay between points c and b. When the switch is closed, no current will flow through the tube; when it is open, plate current will flow during the half cycle that the plate is positive, giving a series of half-cycle pulses through the relay. A condenser may be used, as indicated by the dotted lines, to smooth out these pulses, maintaining a nearly constant current through the winding and thus preventing chattering of the contacts.[1]

[1] Strictly speaking, the presence of the condenser will alter the current flow through the tube, producing short pulses of high current as in the condenser-input filter (see Fig. 7-24, p. 183).

The operation of the circuit of Fig. 8-10 may be more clearly seen with the aid of Fig. 8-11. Curve 1 is a plot of the plate voltage, and the dotted curve 2 is the voltage that will just prevent the flow of plate current during the positive half cycle of plate voltage. Curve 3 is the transformer voltage applied to the grid through resistor R_1 when the switch is closed. It may be seen that this voltage is sufficiently negative to prevent passage of any plate current during the positive half cycle of plate voltage, and, of course, no current will flow during the negative half cycle regardless of the magnitude of the applied positive grid voltage. Thus no plate current flows during the time the switch is closed.

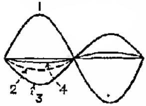


FIG. 8-11.—Relative magnitudes of the voltages of Fig. 8-10.

When the switch is opened, the transformer applies a voltage 4 to the grid through the resistors R_1 and R_2. This voltage is seen

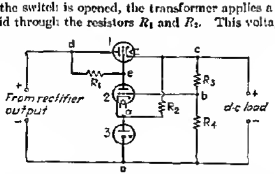


FIG. 8-12.—Circuit of a vacuum-tube voltage regulator.

to be more positive than the cutoff voltage 2, and plate current will flow throughout the positive half cycle of plate voltage.

The resistance R_1 serves the same function as the dotted one in Fig. 8-9. The resistance R_2 limits the flow of grid current during the positive half cycle of grid voltage in the same manner as R of Fig. 8-2.

Voltage-regulated Power Supplies.[1] There are many applica-

[1] For more detailed information on voltage-regulated power supplies, see the following: Alexander B. Bereskin, Voltage-regulated Power Supplies, *Proc. IRE*, **31**, p. 47, February, 1943; W. R. Hill, Jr., Analysis of Voltage-regulator Operation, *Proc. IRE*, **33**, p. 38, January, 1945.

tions using direct current from vacuum-tube rectifiers where special means must be provided to maintain a constant output voltage. One such example is the d-c amplifier described on page 307. Figure 8-12 shows the circuit of a typical regulator which may be connected across the output terminals of any of the rectifiers and filters previously described. With a circuit of this type the voltage may be kept constant to within less than 1 volt with a normal output of a few hundred volts. Since the regulator tends to maintain constant voltage, it will materially reduce any ripple that may remain in the output of the filter supplying the regulator; thus regulated power supplies not only provide a voltage which is independent of variations in supply voltage and in load but which is more thoroughly filtered than that of any reasonably designed unregulated supply.

The operation of the circuit of Fig. 8-12 is as follows. Tube 1 is connected in series with the output and the voltage on its grid is varied in such a manner as to maintain a nearly constant output voltage by changing the drop through the tube whenever there is a change in the supply voltage or in the load demand. The grid voltage of tube 1 is controlled by the plate current of tube 2 since this current determines the drop across R_1. The grid voltage of tube 2 is in turn controlled by the voltage across R_4 which, as is evident from the circuit, is proportional to the output voltage. Tube 3 is a cold-cathode, gas-filled, glow-discharge tube which maintains a constant difference in potential across its terminals as long as current is flowing through it. Resistor R_2 supplies the necessary current to maintain a discharge through the tube.

Example. The reason for using a cold-cathode tube rather than a resistor between points *a* and *c* may best be explained by using a numerical example. Let the output voltage be 275 and let the drop in tube 2 at normal plate current be 100 volts. The grid of tube 1 should be somewhat negative with respect to its cathode, say 25 volts. If the potential of the negative d-c lead is used as reference the potentials of various parts of the circuit will be: point *c*, 275; point *a*, 250; point *a*, 150. If the grid voltage of tube 2 must be 5 volts negative with respect to its cathode to give normal current the potential of point *b* should be 145 volts and the resistors R_3 and R_4 must be so proportioned as to give this voltage with 275 volts output. If the output voltage is now assumed to change to 270, the potential of point *b* will become 142.5 volts. But the cathode potential of tube 2 is maintained at 150 volts by tube 3 so that the voltage between grid and cathode of tube 2 is increased negatively from 5 to 7.5, thereby dropping its plate current, decreasing the current through R_1, and making the grid of tube 1 more positive

(less negative) by an amount that depends upon the gain of tube 2. The drop through tube 1 is thereby reduced and the voltage across the load again rises, approaching the original 275. The output voltage cannot be raised to exactly 275 by this circuit since some drop in voltage is required to cause the regulator to operate; if tube 2 is a high-mu tube, the output may be kept constant to within a small fraction of 1 per cent. Pentode tubes are frequently used because of their high gain, the screen-grid voltage being obtained from a suitable tap on R_2 or R_4.

Now suppose that a resistor had been used between a and o, Fig. 8-12, instead of the cold-cathode tube. A change in output voltage of 5 would then have reduced the potential of point a in the same ratio as that of point b, and the change in voltage between grid and cathode of tube 2 would have been very small. (The change would have been $(150/275)(275-270) - (145/275)(275-270) = 0.091$ volts, where 150 and 145 are the original voltages of the cathode and grid, respectively.) With the cathode voltage maintained constant by the cold-cathode tube the entire voltage change in the output circuit is applied between grid and cathode of tube 2 except for the approximately two to one reduction through R_3 and R_4. (The change would be $(145/275)(275-270) = 2.5$ volts.)

A principal objection to this type of voltage regulator is its low efficiency. The drop through tube 1 is normally of the order of 100 volts or more so that for output voltages of a few hundred volts, for which the regulator is most commonly used, the loss in tube 1 is a considerable portion of the total power supplied to the regulator. On the other hand the total power handled by this type of regulator is usually not more than a few hundred watts, and the loss of energy in the regulator is of small importance considering the excellent voltage regulation obtainable.

It is possible further to refine the circuit of Fig. 8-12 by introducing voltages on the grid of tube 2 which are a function of the input voltage and of the output current. This will provide a compounding action which will further improve the performance of the regulator and may, if desired, be designed to give a rising output potential with increasing load or with decreasing applied voltage.[1]

Regulators may also be designed to produce a constant current instead of a constant voltage. Such units are commonly known as current stabilizers.[2] The circuits of current stabilizers are

[1] See Hill, *loc. cit.*

[2] For suitable circuits and a discussion of the performance of current stabilizers see W. R. Hill, Jr., Analysis of Current-stabilizer Circuits, *Proc. IRE*, **33**, p. 785, November, 1945; J. N. Van Scoyoc and E. H. Schultz, Current Stabilizers, *Proc. IRE*, **32**, p. 415, July, 1944.

fundamentally the same as those of constant-voltage regulators, but the principal excitation for the control tube is obtained from a resistor in series with the output and is therefore a direct function of the output current instead of the potential.

Other Control Applications of the High-vacuum Tube. The control circuits presented in this book are but a few samples of many similar applications of the high-vacuum tube. They are illustrative only, the circuits actually used being subject to many variations in equipment and circuit design to fit a particular need. Even a moderately complete coverage of such circuits and applications is far beyond the scope of this book. Nevertheless it is believed that the information presented herein will enable the reader to trace out and understand most commercial circuits of this type and to devise modifications to meet many of his own problems.

2. GAS-FILLED TUBES

Gas-filled tubes are used widely for control purposes, since they are capable of passing comparatively large currents with very little tube drop. As stated in Chap. 4, any of the various types of gas-filled tubes may be equipped with a grid or other means of control and thus used as a control device. The type most commonly used is the hot-cathode, mercury-vapor tube which, when equipped with a grid, is known as a *thyratron*. The mercury-arc tube may also be equipped with a grid, and the ignitron may be controlled by using a thyratron as the igniting tube to control the instant of ignition.

It was further pointed out in Chap. 4 that the grid in a thyratron (or in a mercury-arc tube) will not control the magnitude of the current flowing through the tube but is capable only of determining the time of ignition. Neither can the grid stop the flow of current once it has started. Thus the control that these tubes can exercise is strictly of the on-off type or, more particularly, of the "on" type, special means being provided to secure an "off" action, as outlined in the succeeding sections. By special means it is also possible to vary the *average* current flowing when alternating current is applied to the tube and load.

D-C Circuits for Use with Thyratrons. A circuit such as Fig. 8-1 is worthless if used with a thyratron tube. It is true that, if the grid potential is sufficiently negative before the plate voltage

is applied, the tube will not conduct current after the plate is energized, but moving the potentiometer toward the positive end will eventually increase the grid potential until the tube begins to conduct, whereupon no further effect on the plate current may be observed as the potentiometer is moved in either direction. As explained in Chap 4, the positive ions produced by ionization of the mercury vapor form a sheath around a negative grid, completely neutralizing its charge.

The circuit of Fig. 8-8 would be equally ineffective, since throwing the switch from position 1 to position 2 would fire the tube, but throwing the switch back would in no way reduce the flow of current.

The plate current of a thyratron with d-c plate supply may be stopped either by opening the plate circuit momentarily (with a sufficiently negative potential simultaneously applied to the grid to prevent the tube from firing again when plate voltage is reapplied) or by momentarily inducing in the plate circuit an emf of opposite polarity to that of the plate supply and of at least equal magnitude The latter method is illustrated by the circuit of Fig 8-13 where the voltage E_{cc} is more negative than the firing voltage of the tube. If the switch is open when plate voltage is applied to the tube, no plate current will flow, but conduction will start immediately upon closing the switch, owing to removal of the negative grid voltage. Opening of the switch will then have no further effect on the plate current, which will continue to flow unimpeded However, if an impulse is applied to transformer T, of sufficient magnitude and correct polarity, the total voltage between plate and cathode may be made momentarily negative, the plate current will cease, and the grid will again assume control. It is necessary that the pulse be of sufficient duration to maintain the plate at a negative potential until the tube has deionized, as reapplication of plate voltage before deionization has been completed will reestablish the arc in the tube regardless of the grid potential. Thyratrons normally have a deionization time of about 100 to 1000 μsec.

The stopping pulse applied to the tube in the circuit of Fig. 8-13 may be secured from the discharge of a condenser, a momentary application of direct current to the primary of the transformer, firing of another thyratron (as in the inverter), or any other suitable source. The most desirable source of this voltage

will depend upon the use to which the circuit is to be put, but it is generally supplied automatically by the device that is under control.

A condenser-discharge method of stopping the flow of plate current in a thyratron is illustrated in Fig. 8-14. When switch S_1 is closed, current flows through the tube setting up a voltage across the load and charging condenser C with the polarity indicated. Reopening of S_1 will have no further effect but if S_2 is closed, the

FIG. 8-13.—Circuit for stopping the flow of plate current in a thyratron by momentarily inducing a high negative potential in series with the plate.

flow of current will cease. Closing of S_2 connects the positive terminal of the condenser to the cathode and, since the negative terminal is connected to the anode of the thyratron, the anode will become momentarily more negative than the cathode, the flow of current through the tube will be interrupted, and the grid will

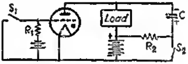


FIG. 8-14.—Circuit for stopping the flow of current in a thyratron by the discharge of a condenser.

again assume control if S_1 is still open. As in the preceding circuit, the negative voltage must remain on the anode for a sufficient length of time to permit deionization of the tube. This requirement, together with the expected load impedance, determines the size of the condenser needed.

A-C Operation. Thyratrons are most commonly used with a-c supply, as in the circuit of Fig. 8-15. A transformer supplies the necessary voltage for the plate and grid, with a separate winding to supply heater power for the cathode. A resistance R_2 is shown in the grid circuit to limit the grid current, positive or negative, to a safe value. If the grid voltage does not greatly exceed the starting potential, this resistance may be omitted.

The wave shapes of the plate and grid voltages are shown in Fig. 8-16. When the grid switch is closed, the grid voltage is at all times more negative than the starting voltage during the positive half cycle of plate voltage. During the negative half cycle of plate voltage, the grid is positive and may draw some current itself, the amount being determined by the voltage applied and the size of the resistance R_2, but no appreciable current can flow to the plate during either half cycle.

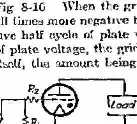


FIG. 8-15.—Thyratron control circuit with a-c supply.

Opening the switch raises the potential on the grid during the half cycle that the plate is positive and permits a flow of current through the tube. At the end of this half cycle the plate becomes negative and will not attract electrons and the flow of current ceases, but each time the plate again becomes positive, the flow of current resumes. If the switch is again closed at any time during the cycle, no change in the cycle of the plate current will result until the plate voltage again changes from negative to positive. At that point, the flow of plate current that would start with the switch open will fail to start with the switch closed, and the current will remain at zero. By this means the grid is apparently able to stop the flow of current through the tube as well as to start it, although in the strict sense of the word it does not actually stop the flow but simply prevents its starting again after the plate has done the stopping.

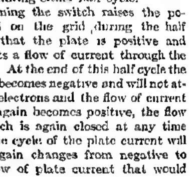


FIG. 8-16.—Relative magnitudes of the voltages in the circuit of Fig. 8-15.

The flow of plate current through this tube is evidently unidirectional but pulsating, having the general appearance of Fig. 8-17. The average value, assuming a half sine-wave shape, will be 0.318 times the crest value.

It is of interest to note from Fig. 8-16 that the minimum alternating potential which will prevent the tube from firing is determined by the potential at point x, not by that at point y as might

at first be supposed. This is because the curve of starting voltage is not sinusoidal, so that if the switch in Fig. 8-15 is closed and the voltage between a and c in that figure is gradually reduced in some manner (plate voltage remaining constant), the tube will fire as soon as curve a, Fig. 8-16, intersects the curve of starting voltage, at point x.

FIG. 8-17.—Wave shape of the thyratron current in the circuit of Fig. 8-15.

The starting voltage curve of a given tube may be determined from the characteristic curves, such as those of Fig. 4-25, by first reading the plate supply voltage at any given instant of time from curve d, Fig. 8-16, and then finding the corresponding grid voltage from the appropriate curve of Fig. 4-25. Such a procedure will give the dotted curve of Fig. 8-16.

On the other hand plotting of the starting voltage curve is not needed to design the transformer of Fig. 8-15. To illustrate, one of the curves of Fig. 4-25 is reproduced as curve 1 in Fig. 8-18. We may also draw a straight line to represent the corresponding instantaneous grid and plate voltages supplied at points a and d in the transformer. If the slope of this curve is adjusted to produce tangency with curve 1, it will represent the minimum ratio of grid-to-plate voltage (and, therefore, the minimum grid voltage for a given plate voltage) which will prevent the tube from firing. If the rms value of the plate voltage to be supplied by the transformer is known, the corresponding minimum value of rms grid voltage which must be supplied at point a may be read from curve 2 of Fig. 8-18.

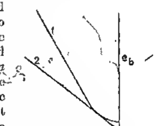


FIG. 8-18.—Showing how to determine the cutoff voltage of a thyratron.

Plate-current Control by Grid Phase Shift. The average value of the current flowing through a thyratron tube with a-c supply may be varied at will from zero to a maximum by shifting the phase of the grid voltage relative to the plate. A circuit com-

monly used for this purpose is shown in Fig. 8-19. The plate is energized in the usual manner, but the grid voltage is supplied from a coil located in a three-phase revolving field. By rotating the inner coil, any phase position of the grid voltage relative to that of the plate may be obtained.

Grid phase shift may also be obtained from a single-phase supply by means of the circuit of Fig. 8-20. Here the grid is energized

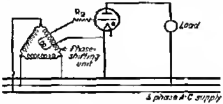


FIG. 8-19.—One method of securing grid phase shift with a three-phase supply.

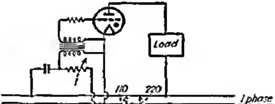


FIG. 8-20.—A method of securing grid phase shift with single-phase supply.

from the secondary of a transformer the primary of which is connected between the neutral wire of a three-wire, single-phase system and the mid-point of a series condenser and resistance, bridged across the 220-volt supply. Varying the resistance will rotate the grid phase through 180 deg, with a magnitude that would be constant except for the effect of the exciting current of the transformer and of any grid current in the tube. It is important that the transformer have a high no-load reactance (low magnetizing current), otherwise resonance between this reactance and that of the condenser may cause undesirable effects.

If the circuit of Fig. 8-20 is placed in operation, the grid may be found to have no effect on the magnitude of the current flowing, except to cause it to drop abruptly from maximum to zero when

the resistance has been nearly all removed. The reason for this action is explained in the next section, but the remedy is to interchange the connections on the secondary of the transformer so as to reverse the polarity of the grid voltage.

The adjustable resistance of Fig. 8-20 may be replaced by a high-vacuum tube, the grid voltage of which is varied to control the current through the thyratron. This permits variation of the thyratron current with the expenditure of a very small amount of

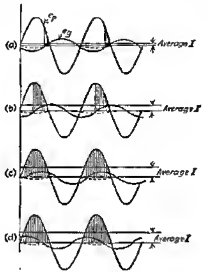


FIG. 8-21.—Wave shapes of currents and emfs in a thyratron with grid phase shift.

power and thus permits control by very sensitive devices of circuits carrying large amounts of power.[1]

Wave Shapes. The effect of phase shift of the grid is illustrated in Fig. 8-21 for four different positions of the phase-shifting coil of Fig. 8-19. In curve *a* the grid is nearly 180 deg out of phase with the plate, and current does not start to flow until almost the end of the positive half cycle of the plate voltage, as indicated by

[1] For detailed circuits and discussion of such methods of control, see Samuel C. Coroniti, Analysis and Characteristics of Vacuum-tube Thyratron Phase-control Circuits, *Proc. IRE*, **31**, p. 653, December, 1943.

the shaded area under the plate-voltage curve. The dotted curve, representing the starting voltage, is seen to be more positive than the grid voltage throughout the major portion of the half cycle, and the plate current remains zero until the two voltages become equal

Curves *b* and *c* represent two other positions of the phase-shifting coil in which the phase of the grid voltage becomes increasingly closer to that of the plate. The period throughout which the current flows is seen to be greater as the grid comes more nearly into phase with the plate Shifting the phase of the grid evidently does not increase the *instantaneous* value of the current during the time it is flowing but increases the average value by varying the *width* of each *pulse*

If a thyratron is passing maximum current through a given load owing to an in-phase relationship between grid and plate voltages, a shift in phase of the grid in the *correct* direction will cause a decrease in average current as illustrated, but if the shift is made in the opposite direction, the plate current is unaffected This is illustrated in curve *d* (Fig 8-21), where the phase shift away from the in-phase position is the same in amount as for curve *c* but in the opposite direction. At the beginning of the positive half cycle of the plate voltage the instantaneous grid voltage exceeds the starting voltage, and plate current will flow. It does not matter that the grid voltage becomes negative during the latter part of the positive half cycle of plate voltage, since the positive ions already present will nullify the effect of the negative grid Regardless of how far the phase of the grid may be shifted to the left, the plate current will continue to flow at full strength until such time as the grid becomes out of phase by 180 deg, when the current will drop abruptly to zero.

Circuits with Alternating Current through the Load. In the thyratron circuits presented thus far the load has been inserted directly in series with the tube and thus received a pulsating direct current. Very often it is desirable that the load current be a true alternating one. One method of accomplishing this was shown in Fig 8-6 for high-vacuum tubes which may be applied to thyratrons by merely replacing the potentiometer with some type of phase shifter. The current flowing in the load will then be alternating but not sinusoidal, since the pulses of current

passed by thyratrons are not half sine waves except when maximum current is flowing (see Fig. 8-21).

A somewhat more desirable method is to use saturable reactors, since the reactor not only permits a nearly sinusoidal flow of current through the load but also serves as an amplifier of the tube current. The circuit of Fig. 8-22 illustrates the use of two tubes in this manner.[1] Any type of grid phase shifter may be used; the one shown here operates from a single-phase circuit by connecting two coils, mounted in space quadrature, across a condenser C and resistance R to provide two fields in time and space quadrature. The grid pickup coil G may be rotated to secure the desired phase.

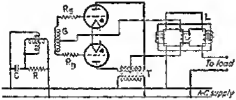


FIG. 8-22.—Circuit of two thyratrons, with grid phase shift, controlling the current through a load circuit by means of a saturating reactor.

The saturable reactor L has three windings, two in series with the load and the other in the common plate circuit of the two thyratrons. When the thyratron current is zero, the reactor impedance is maximum, reducing the load current to a low value. As the thyratron current increases, the core of the reactor will saturate (since the thyratron current contains a large d-c component), and the impedance will drop. With proper reactor design, a continuously variable control may be obtained over a considerable range in load current.

The two windings in series with the load are so connected that they induce zero voltage into the third winding. This is important, as otherwise the actual voltage applied to the plate of each thyratron would be the sum or difference of the voltage of the transformer T and the voltage induced in the saturating winding

[1] Two tubes should always be used, in a full-wave circuit, since the reactor tends to keep constant current flowing. A single tube would be equivalent to the use of an L-section filter with a single-phase, half-wave rectifier (see Chap. 7, p. 184).

on the reactor L. If but one load winding was used, the magnitude of the voltage induced in the reactor might easily be sufficient seriously to affect the performance of the thyratrons, since the saturating winding is often wound with many turns to enable a small tube current to produce saturation.

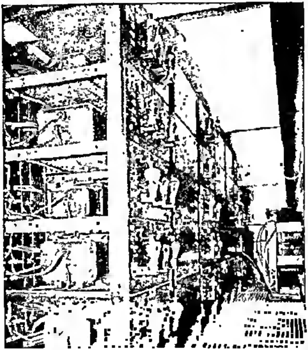


FIG. 8-23.—Tube panels, reactor racks, and reactors of a thyratron saturating reactor control, for dimming theater lights. (*Courtesy of General Electric Company.*)

Figure 8-23 shows an installation of several such units for dimming theater lights. The reactors may be seen behind the panel.

Controlled Rectifier.[1] Since the thyratron will pass current in one direction only, it can be used as a rectifier. The presence

[1] C. C. Herskind, Grid-controlled Rectifiers and Inverters, *Elec. Eng.*, **53**, p. 926, June, 1934; also Caldwell B. Foos, Vacuum-tube-controlled Rectifier, *Elec. Eng.*, **53**, p. 568, April, 1934.

of the grid adds nothing to its rectifying properties but provides a means of varying the average output voltage, a feature that is frequently of the utmost value. A typical circuit of this type is shown in Fig. 8-24 using two thyratrons as a full-wave, single-phase rectifier.[1] Transformer T_1 supplies the plate voltage to the two tubes; T_2, the cathode heating power; and T_3, the variable phase excitation to the grids. This voltage is applied to the grids through the center-tapped transformer T_4 in order that the grid and plate in each tube shall bear the same relative phase to each other. By means of the phase-shifting transformer T_3 the *average* current through the tubes may be adjusted to any desired

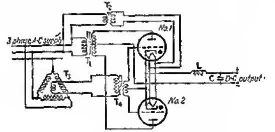


FIG. 8-24.—Circuit of a grid-controlled, single-phase, full-wave rectifier.

value. The output voltage is, of course, a direct function of the current flowing through the load.

The L-section filter shown in Fig. 8-24 will require a continuous flow of current if the choke is equal to or larger than its critical minimum value (as it should be).[2] As the grid phase is shifted, each tube will fire at a later point in the cycle but will then *continue to conduct for 180 deg* (at which time the other tube fires). Thus the effect is to shift the phase position of the conducting period for each tube but without changing its length, as shown in Fig. 8-25.

The curves of Fig. 8-25 depict the theoretical voltage relations for four different firing angles assuming that the inductance is in

[1] Except for the method of grid control (and any type of grid phase-shifting device may be used), this circuit is similar to that of Fig. 8-22. Consequently the thyratrons used for control purposes through saturable reactors actually constitute a full-wave rectifier.

[2] See the discussion on p. 190, for the meaning of the minimum value of inductance.

each case at least equal to its critical value. The voltages supplied to the anodes of the tubes by the two halves of the transformer T_1 are shown as sine waves of opposite polarity (the midpoint of the transformer being the reference, or zero potential, point). If the tube drop is neglected, the potential of the cathodes of both tubes must be equal to the potential of that anode which is

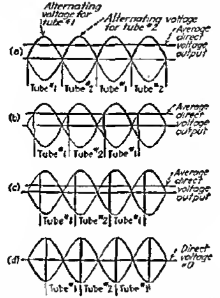


FIG. 8-25.—Curves showing the effect of grid phase shift on the output voltage of the circuit of Fig. 8-24.

conducting at the moment, and the cathode potential will be as indicated by the heavy lines. Inspection of the circuit of Fig. 8-24 will show that this is also the potential impressed on the input to the filter.

Figure 8-25a shows the conditions for zero phase shift. Each tube conducts throughout one entire half cycle (neglecting the small effect of tube drop), and the average output voltage is equal to the average voltage of one-half the transformer T_1, less the drop in tubes and choke.

If the grid phase is now shifted to give a phase shift angle $\theta = 30$ deg, the point of firing will be shifted approximately 30

deg beyond the point of zero impressed voltage. The tube that was conducting during the previous half cycle must then continue to conduct until the other tube fires or for 30 deg after its supply voltage has become negative, since the current through the choke cannot become zero under the assumptions made. It is enabled to do so by the voltage induced in the choke which is of such magnitude as to apply a net positive voltage between its anode and cathode. This is illustrated in the (*b*) set of curves in Fig. 8-25. The average voltage (which is of course the direct output voltage, neglecting tube and choke drop) is evidently lower in this case than when $\theta = 0$ deg, and thus the output voltage is controllable by the action of the grid.

The other two sets of curves in Fig. 8-25 depict the operation for $\theta = 60$ deg and $\theta = 90$ deg. It is seen that zero output voltage is obtainable with a phase shift of only 90 deg (not 180 deg), but it should be remembered that these curves were drawn by assuming the inductance to be at least equal to its critical value which, for zero load current, would be infinite. Thus the output voltage and current will be brought to nearly zero with a phase shift of 90 deg, but absolute zero can be secured only with a phase shift of 180 deg, as described on page 234 where no inductance whatsoever was considered.[1]

This method of controlling the voltage of a rectifier may, of course, be applied to polyphase rectifiers as well as to single-phase units. It is necessary only to supply the grids of the various tubes in the same relative phase relations as their anodes and then apply phase shift to all grids simultaneously.[2]

Inverters. Thyratrons may be used to convert direct current to alternating current and when so used are commonly known as *inverters*. There are two distinct types of thyratron inverters: (1) those which are separately excited and (2) those which are self-excited. Separately excited inverters are of the nature of amplifiers, as they require a small amount of a-c power for their excitation but are capable of producing a much larger amount of a-c power in the output circuit. Self-excited inverters are more

[1] An excellent discussion of the effect of the inductance on the performance of this type of rectifier together with curves of the critical value of inductance to be used is given by W. P. Overbeck, Critical Inductance and Control Rectifiers, *Proc. IRE*, **27**, p. 655, October, 1939.

[2] A typical circuit is given by S. R. Durand and O. Keller, Grid Control of Radio Rectifiers, *Proc. IRE*, **25**, p. 570, May, 1937.

of the nature of oscillators, as they supply their own excitation losses from the a-c output of the circuit.

Separately Excited Inverters. A typical circuit of a separately excited inverter is shown in Fig. 8-26. All the energy taken from the output terminals is drawn from the d-c power supply indicated in the lower part of the figure. The a-c excitation must be supplied from an available low-power source such as a vacuum-tube oscillator,[1] a mechanical vibrator, or other similar source. The cathodes may be heated from the d-c power supply through a series resistance R_2 as shown, or they may be heated from this source only while the inverter is being put into service, drawing

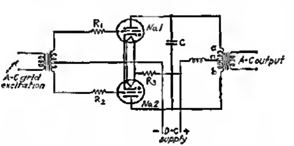


FIG. 8-26.—Circuit of a separately excited thyratron inverter, parallel type.

power from a transformer off the a-c output after the inverter is in operation.

The operation of this inverter may be studied by first considering the half cycle of excitation voltage during which the grid of tube 1 is positive and assuming that tube 2 is not drawing plate current. Plate current begins to flow in tube 1 as soon as the grid becomes more positive than the starting voltage, and as this current increases, an emf will be induced in the output transformer which will be impressed across the condenser C causing it to charge with a negative polarity toward the plate of tube 1. An emf will also be induced into the secondary of the output transformer and so be impressed on the load.

At the start of the next half cycle the grid of tube 1 will become negative, but this will, of course, have no effect upon the flow of plate current, as was pointed out in Chap. 4. However, the grid

[1] See Chap. 11 for a discussion of vacuum-tube oscillators

of tube 2 will become positive, and this tube will fire, causing current to flow from the d-c source through the lower half of the output transformer. The voltage drop between anode and cathode of tube 2 is, of course, very low during conduction, so that the lower terminal of condenser C will be virtually connected to the negative terminal of the d-c supply. Since it had been previously charged by tube 1 with negative polarity on its upper plate, the full negative charge of this condenser will momentarily be applied between the plate and cathode of tube 1, permitting its grid to resume control. Of course, if the grid of tube 1 were not negative at the time this action takes place, both tubes would conduct current after condenser C discharged, thus short-circuiting the d-c supply.

The action described in the preceding paragraph continues to repeat itself, the action of the two tubes being interchanged at the beginning of successive half cycles.

Condenser C is known as a *commutating* condenser because it is provided solely for the purpose of commutating the direct current from one tube to the other at the end of each half cycle of the exciting voltage. Its size must be such as to provide a negative charge of sufficient duration to enable the grid to assume control, usually of the order of 1 to 5 μf.

A leading power-factor load will obviously serve the same purpose as the commutating condenser C. On the other hand, a lagging power-factor load will tend to reduce the effect of C, so that much larger condensers must be used in inverters required to handle inductive loads. Thus it is important to know the type of load that is to be supplied before designing an inverter.

Since the wave shape of the plate current flowing through the primary of the transformer is far from sinusoidal, the output of this inverter might be expected to contain many harmonics. Such is found to be the case, although the wave shape is surprisingly good considering the nature of the action producing the a-c output.

Tompkins[1] gives the operating curves of this circuit, and they are reproduced in Fig. 8-27. The efficiency is seen to be quite high which is characteristic of the gas-filled tubes.

[1] Frederick N. Tompkins, A Parallel-type Inverter, *Trans. AIEE*, 51, p. 707, 1932; also published in condensed form in *Elec. Eng.*, 52, p. 253, April, 1933.

The two thyratrons may be operated in a series type of circuit as compared with the parallel arrangement of Fig. 8-26. A typical circuit of this type is that of Fig 8-28. Consider that tube 1 is conducting and that tube 2 is idle. Current flows from the positive terminal through the primary of the output transformer T, the upper half of the choke L, and tube 1, charging condenser C, with positive polarity on its left terminal. As the polarity of the grid excitation voltage changes, tube 2 will fire, causing the condenser C to discharge, thereby inducing a voltage of opposite polarity in the output transformer and inducing a high negative voltage between terminals a and n of the choke L, sufficient to extinguish the arc in tube 1 and permit its negative grid to assume control again

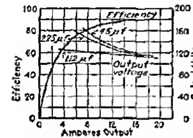


FIG 8 27.—Operating characteristics for the circuit of Fig 8-26 The capacitances refer to the commutating condenser

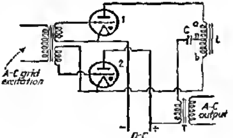


FIG 8 28.—Circuit of a separately excited thyratron inverter, series type.

On the next half cycle of the excitation voltage tube 1 will again fire, causing the condenser to charge. A large voltage is again built up across the inductance L, and a positive voltage appears between b and n which is applied to the cathode of tube 2 Since this will have the same effect as applying a negative voltage to the plate, plate current will cease to flow and the grid can again

assume control. This cycle is repeated as the polarity of the excitation voltage reverses thus producing an a-c output across the secondary of transformer T of the same frequency as the excitation voltage but of much higher power.

Since the same polarity of voltage across L is required to extinguish either tube 1 or tube 2, it might seem that both tubes should cease to conduct simultaneously. That they do not is due to the difference in the charge on condenser C, this condenser being fully charged when tube 2 fires and at a minimum charge when tube 1 fires. The potential applied to the plate of tube 1 is

$$E_{b1} = E_0 - E_c - E_t$$

where E_0 = supply voltage

E_c = voltage across condenser C

E_t = voltage across half coil L plus voltage across T

while the plate voltage of tube 2 is

$$E_{b2} = E_c - E_t$$

When tube 2 fires, E_c is very nearly equal to E_0 and E_t is somewhat less than E_c. The above equations therefore show that tube 1 will have a negative plate voltage while that of tube 2 is somewhat positive. When tube 1 fires, condenser C will have discharged and its voltage will be quite small; thus the equations show that the potential applied at that time to tube 1 will be somewhat positive while that applied to tube 2 will be negative. These are the necessary conditions for commutation of the current from one tube to the other.

Series inverters are generally preferred to the parallel type owing to the reduced probability of a short circuit of the d-c supply through the failure of a tube to commutate. This improvement is due to maintenance of negative voltage on the plate of a thyratron at the end of its conducting period for a longer period of time in the series type than in the parallel type, thus increasing the time for the tube to deionize. With suitable design this period may be made to approach 90 deg.[1] Suitable protective devices should of course be provided in the d-c line with either type of inverter to open the circuit in case of commutation failure.

Self-excited Inverters. Either the parallel or the series type of

[1] C. A. Sabbah, Series-parallel Type Static Converters, *Gen. Elec. Rev.*, 34, p. 288, May, 1931.

inverter may be made self-excited by supplying the grid excitation from the a-c output terminals through suitable circuits. Figure 8-29 shows such an application to the parallel type of inverter. T_1 is the output transformer, and C_1 is the commutating condenser, as in the separately excited inverter. The grids of the two tubes are excited by means of transformer T_2 which is energized from the output transformer as shown.

When the d-c circuit is first closed to start the inverter, both tubes would start conducting at once if a means were not provided to discriminate against one or the other. Condenser C_3 performs this function, being charged through R_1 and L_1, with the return circuit through R_3 and half of transformer T_2. The charging current of this condenser induces a positive voltage in that side of transformer T_2 supplying the grid of tube 1 and consequently

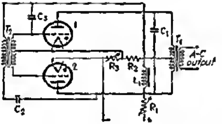


FIG. 8-29.—Circuit of a self-excited thyratron inverter, parallel type.

induces a negative voltage on the grid of tube 2. Therefore, tube 1 will be the first to conduct, and tube 2 will be held inoperative.

At the same time the plate current drawn by tube 1 through the upper half of transformer T_1 will induce a negative voltage on the upper terminal of that half. The lower half of the transformer will therefore impress a positive voltage on the plate of tube 2 which is a function of the charging period of condenser C_1. Eventually tube 2 will become conducting as the induced voltage in transformer T_2 falls off at the end of the charging period of condenser C_1 and allows the grid to become positive. When this occurs, condenser C_1 will apply a high negative voltage to the plate of tube 1, just as described for the separately excited inverter, to stop the flow of its plate current.

The impedance of transformer T_1 prevents the instantaneous discharge of condenser C_1 thus maintaining its negative voltage

sufficiently long for the simultaneously induced voltage in transformer T_2 to swing the grid of tube 1 negative and prevent reestablishment of current through this tube for another half cycle.

The sequence of events just described continues to occur with the action of the two tubes interchanged, the frequency of the output being determined primarily by the constants of the output circuit.

A series type of inverter is shown in Fig. 8-30. If tube 1 is conducting, current will flow from the d-c source, charging condenser C_2. The current also passes through the resistance R_2 and so maintains a negative voltage on the grid of tube 2 which prevents that tube from firing. As condenser C_2 becomes charged, the current flow ceases, causing tube 1 to stop conducting and removing the negative voltage from the grid of tube 2. Firing of tube 2 follows, and condenser C_2 then discharges, setting up a negative voltage across R_1 which is applied to the grid of tube 1. Also, the drop across coil L_1 applies a negative voltage to the anode of tube 1 through the condenser C_1, which has zero voltage across its terminals at the time when tube 2 fires, since it was discharged by tube 1 while C_2 was charging. The negative potential remains on the anode for a sufficient length of time to permit the tube to deionize and so permit the grid to resume control. After condenser C_2 has discharged and C_1 charged, the current flow through tube 2 ceases and tube 1 fires. In this case the discharge of condenser C_1 sets up a voltage across L_1 of the opposite polarity, thus applying a negative voltage to the anode of tube 2 which permits it to deionize. This cycle is then repeated at a frequency determined by the circuit constants.

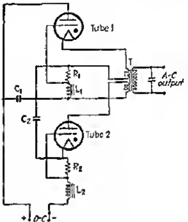


FIG. 8-30.—Circuit of a self-excited thyratron inverter, series type.

Mercury-arc Inverters.[1] Mercury-arc rectifiers may be equipped

[1] For more details, see H. D. Brown, Grid-controlled Mercury-arc Rectifiers, *Gen. Elec. Rev.*, 35, p. 439, August, 1932.

with grids and used in any of the applications previously listed for thyratron tubes, including inverters. Their performance is very similar to that of thyratrons except for their greater current-carrying capacity. However, they are being largely superseded by ignitrons, and no further discussion of mercury-arc tubes will be presented here.

Phase-shift Control of Ignitrons. The current flow through ignitrons may be controlled just as readily as that through thyratrons by means of a small thyratron in the igniter circuit. The

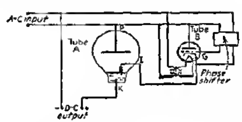


FIG. 8-31.—Circuit of an ignitron with grid phase-shift control on the igniter tube.

latter tube is controlled by any of the methods previously described for thyratrons, generally a phase-shifting method. A circuit for performing this control is shown in Fig. 8-31. It is seen to be similar to that of the ordinary ignitron rectifier of Fig 7-43 except for the use of a grid-controlled rectifier for tube B together with the necessary associated circuits for shifting the phase of its grid excitation. The grid control on tube B determines the time in the cycle at which the igniter is energized and therefore determines the point at which the ignitron starts to draw current.

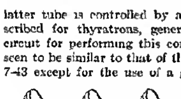


FIG. 8-32.—(a) Ignitron anode current and (b) igniter current for an ignitron with a small phase shift on the grid of the igniter rectifier.

This process is further illustrated by the curves of Fig. 8-32 for a small phase shift, where *a* shows the current flow through the ignitron and *b* shows that through the igniter circuit. It may be seen that the flow of current through the latter circuit is essentially the same as that without grid control, as shown in Fig

7-44, except that it occurs at a later point in the cycle. The ignitron current also starts at this later point in the cycle and then continues to flow until the anode voltage drops to zero. Evidently the current flow through the anode circuit is exactly like that flowing in the anode circuit of a thyratron with grid phase shift, Fig. 8-21, although the controlling electrode is entirely outside the tube.

The ignitron may be used with a control circuit of this type to perform exactly the same functions as have already been discussed in connection with the thyratron, *e.g.*, the controlled rectifier. It is generally able to handle heavier overloads than the thyratron but requires the use of more auxiliary equipment for its operation.

Ignitron Inverters. Ignitrons may also be operated as inverters. The circuits are fundamentally the same as those for the thyratron except that the excitation voltages must necessarily be applied to the grids of the igniter tubes. Ignitron inverters can be built with much greater power capacity than can thyratron inverters, thus opening up tremendous potential applications in the electric-power field.

Applications of Thyratrons and Ignitrons. The number of applications to which these tubes may be put is virtually unlimited. A few examples are controlled rectifiers, the thyratron motor (an adjustable-speed a-c motor),[1] d-c power transmission,[2] frequency changer,[3] voltage regulator,[4] lighting control,[5] welding control,[6]

[1] E. F. W. Alexanderson and A. H. Mittag, The Thyratron Motor, *Elec. Eng.*, 53, p. 1517, November, 1934; and E. F. W. Alexanderson, M. A. Edwards, and C. H. Willis, Electronic Speed Control of Motors, *Trans. AIEE*, 57, p. 343, June, 1938.

[2] C. H. Willis, B. D. Bedford, and F. R. Elder, Constant-current D-c Transmission, *Elec. Eng.*, 54, p. 102, 1935; C. C. Herskind, New Types of Transformers, *Trans. AIEE*, 56, p. 1372, November, 1937.

[3] A. Schmidt, Jr., and R. C. Griffith, A Static Thermionic Tube Frequency Changer, *Elec. Eng.*, 54, 1063, 1935.

[4] Palmer H. Craig and Frank E. Sanford, An Electronic Voltage Regulator, *Elec. Eng.*, 54, p. 166, 1935.

[5] E. D. Schneider, Thyratron Reactor Lighting Control, *Trans. AIEE*, 57, p. 328, June, 1938. A considerable bibliography of other articles on lighting control is included with Schneider's paper.

[6] T. S. Gray and J. Breyer, Jr., Electronic Control for Resistance Welders, *Trans. AIEE*, 58, p. 361, July, 1939; J. W. Dawson, New Developments in Ignitron Welding Control, *Trans. AIEE*, 55, p. 1371, December, 1936.

speed control of ordinary d-c motors,[1] and many others.[2] A very brief picture of some of these applications follows.

Controlled rectifiers using thyratrons were discussed on page 237. The circuits used with ignitrons differ essentially only in the addition of firing circuits in place of the thyratron grid control.

The thyratron motor may be considered as essentially a d-c series motor without a commutator but with thyratrons providing the commutating action by switching the current from one to another of perhaps six coils. The d-c power for the motor is supplied from regular a-c power sources by a set of thyratrons in a grid-controlled rectifier, although in actual practice the same tubes perform the functions of both rectification and commutation. The motor terminal voltage and, therefore, its speed are controlled by varying the phase of the grid voltage on the thyratrons. An adjustable-speed a-c motor with a remarkably high efficiency is the result.

Direct-current transmission is made possible by rectifying the output of high-voltage transformers at the generator end of a line and then converting back to alternating current at the receiver end by inverters. Direct-current transmission has certain inherent advantages over a-c which make it extremely attractive if tubes can be developed to provide the very high voltages necessary for efficient transmission, together with sufficient power capacity to handle commercial installations.

Lighting control is provided largely through the medium of saturable core reactors, as in Fig. 8-22. The tube used to control the saturation of the core may even be a high-vacuum type where the controlled lighting load is not too great. Such tubes offer the advantage of potentiometer control of their grids rather than the more complex phase-shifting method required for thyratrons.

The use of vacuum tubes in the control of resistance welding is now almost standard practice. With thyratron and, more recently, ignitron control, the flow of welding current may be limited to but one half cycle or even a portion of one half cycle, or it may be extended over a given number of cycles, thus ac-

[1] G. W. Garman, Thyratron Control of D-c Motors, *Trans. AIEE*, 57, p. 335, June, 1938.

[2] D. E. Chambers, Applications of Electron Tubes in Industry, *Elec. Eng.*, 54, p. 82, 1935, also J. W. Horton, Use of Vacuum Tubes in Measurements, *Elec. Eng.*, 54, p. 93, 1935, for a bibliography of vacuum tubes in measurement work.

curately determining the intensity of the heat and the extent to which it spreads.

Operation of Cold-cathode, Grid-controlled Tubes. The cold-cathode, grid-controlled tubes described on page 126 find their most common applications in the control of relays and other low-current loads where continuous stand-by service is required and especially where the actual time of use is low. Since no local power source is required for cathode heating, no power is consumed during idle periods, whereas hot-cathode tubes would draw heating power continuously even though used but a small portion of the time.

A circuit employing the OA4-G tube is shown in Fig. 8-33. The plate of the tube is connected to one side of the 60-cycle line through a suitable relay. The starter electrode is energized from the same source through a voltage divider which supplies a poten-

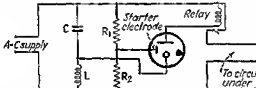


FIG. 8-33.—Circuit for using the grid-controlled, cold-cathode tube of Fig. 4-30 to operate a relay by remote control.

tial slightly less than the starter breakdown voltage, the resistance of the divider being sufficient to prevent the flow of excessive current through the starter when the tube is fired. The cathode is energized from the mid-point of a series condenser and inductance. The tube is fired from a remote point by applying to the power line at that point a current having a frequency of some tens or even hundreds of thousands of cycles. The inductance and capacitance that supply the cathode are resonant to this frequency and sufficient voltage is built up across the inductance to raise the potential between cathode and starter above the breakdown point. The tube therefore will not conduct, and the relay will be deenergized, until the r-f current is supplied from the remote point, after which the tube will break down and close the relay. Upon removal of the r-f source the tube will cease to conduct on the first negative half cycle of voltage supply to its anode. Thus the relay may be controlled remotely over the regular 60-cycle power wires.

The distances over which this control may be operated depend upon the nature of the a-c circuit, whether or not there are transformers interposed between the control point and the tube, the amount of load on the a-c circuit (*i.e.*, the impedance between the wires), etc. When used on 110-volt circuits in buildings, however, it has been found that the normal line impedance is so low that the amount of load has very little effect and that control may be carried on from most points within a given building but not from points outside the building.[1]

The oscillator that supplies the r-f power may be energized directly from the a-c line without a rectifier, whence it will supply r-f power only on the positive half cycles of the a-c line.[2] If these half cycles correspond to those of positive voltage on the anode of the cold-cathode tube, the relay will be operated, whereas if they correspond to the negative half cycles on the tube, the relay will remain deenergized. If another cold-cathode tube is used in a circuit exactly like that of Fig. 8-33 but with the a-c line terminals reversed, it is possible to operate either of the two relays by merely reversing the polarity of the alternating current supplied to the oscillator at the control station.[3]

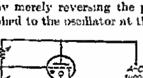


FIG. 8-34.—Circuit using the cold-cathode tube of Fig. 4-27 to operate a relay.

Another application of cold-cathode, grid-controlled tubes is illustrated in Fig. 8-34, the circuit being particularly adapted to the tube of Fig. 4-27. Adjustment of the series condenser (or of the resistor) of Fig. 8-34 will make the tube conducting or nonconducting. Either the resistor or the condenser is normally included in the main unit with the tube and may be adjusted to give the proper sensitivity to the device, but the other is generally a portion of the apparatus or circuit that is to be controlled. For example, the plates of the condenser in Fig. 8-34 may be placed on either side of a chute to indicate the presence of a metallic object, or a burglar alarm may be arranged by extending a wire around the region to be protected, using it as one plate of the con-

[1] Charles N. Kimball, A New System of Remote Control, *RCA Rev.*, 2, p. 303, January, 1938.

[2] See Chap. 11 for oscillator circuits.

[3] Kimball, *loc. cit.* Also A. V. Eastman, Patent No. 1834771.

However, R_g should not be increased to the point where it is comparable to the resistance of the phototube during its unilluminated state, as then an appreciable positive voltage will be applied to the grid even when no light is impinging on the cell, nor should it exceed the grid-cathode resistance of the amplifier tube. A resistance of 5 or 10 megohms is usually satisfactory for gas-filled phototubes, although somewhat higher resistances may be used with high-vacuum phototubes, provided the input resistance of the amplifier tube is made high by keeping its plate voltage low.[1]

Additional stages of amplification may be added if necessary; but if the relay is to be operated when light strikes the phototube and remain unoperated when there is no light, such amplifiers must be of the d-c type (see page 307); *i.e.*, the current in the relay must continue to flow as long as light strikes the phototube. The use of d-c amplifiers is to be avoided since variations in the direct supply voltages of the early stages of the amplifier, even though extremely small, will cause large variations in the plate current of the final amplifier tube. This problem may be solved by causing the light to vary periodically, thus producing an alternating current in the grid resistor of the first tube. The output of this tube may then be amplified by the usual a-c amplifier circuits described in Chap. 9. Pulsations in light may be produced by periodically interrupting the light beam with a mechanical chopper or by the use of a gaseous lamp wherein the light is extinguished twice per a-c cycle, thus giving off a light that pulsates at twice the frequency of the alternating current used to energize the lamp. It is also possible to use a constant light beam and insert a chopper in the grid or plate circuit of the first tube, using either a mechanical device, such as a rotating disk with alternate conducting and nonconducting segments, or an electric discharge type of circuit.

Phototube Control of Thyratrons. A circuit for operating a screen-grid thyratron with a phototube using a-c supply is shown in Fig. 8-36.[2] A voltage divider across the a-c supply provides the thyratron plate voltage between points b and d. The grid

[1] This reduces the probability of ionization of any traces of gas in the tube. Ionization, if present, would cause a small grid current, owing to the flow of positive ions to the negative grid and would therefore lower the input resistance. If a pentode tube is used as the amplifier, its screen-grid voltage must also be low.

[2] See p. 120 for discussion of screen-grid thyratrons.

voltage is the difference between the alternating bias voltage across the resistance cd and the voltage drop across resistance R_g produced by current flowing through the phototube. Plate current will, therefore, flow through the relay on the positive half cycles of the supply whenever there is sufficient light on the phototube to produce a net voltage on the grid that is more positive than the starting voltage.

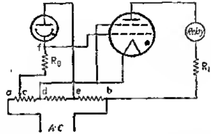


FIG. 8-36.—Circuit showing a thyratron used to operate a relay under control of a phototube.

The voltages applied to the thyratron of Fig. 8-36 are shown in Fig. 8-37, where curve 1 represents the plate voltage and curve 2 the starting voltage; *i.e.*, whenever the grid voltage is more positive than curve 2, the tube will conduct. Curve 3 represents the voltage between points c and d, which is the voltage applied to the grid of the thyratron when no light is impressed on the phototube. As may be seen, it is at all times more negative than the starting voltage, and the thyratron does not conduct. When light strikes the phototube, current flows during the positive half cycle of impressed emf, and the net voltage applied to the thyratron grid is reduced to that of curve 4.[1] It may be seen that this

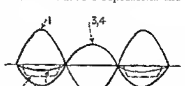


FIG. 8-37.—Relative magnitudes of the voltages in the circuit of Fig. 8-36.

[1] Although it is not a matter of particular importance, it is of interest to note that the bias on the thyratron is unaffected during the negative half cycle of plate voltage, since the phototube also has negative voltage on its plate and therefore will not conduct during that half cycle.

voltage exceeds the starting voltage, and the relay will be operated.

Either high-vacuum or gas-filled phototubes may be used in the circuit of Fig. 8-36, but the latter generally will produce a higher response, especially as it is difficult to use exceptionally high values of R_g with thyratron tubes

Photovoltaic Cells as Control Devices. As shown in Chap. 5 it is possible to use photovoltaic cells to operate a relay *directly* and so avoid the need for vacuum-tube amplifiers. Elimination of the amplifier tube materially reduces both the first cost and the maintenance of the equipment, but the intensity of light required is much greater than when photoelectric tubes are used with suitable vacuum-tube amplifiers Whether a photovoltaic or photoemissive type of unit is to be used must, therefore, be determined by the circumstances surrounding a particular installation.

Phototube (and, to a lesser extent, photovoltaic) control circuits of the type described in the preceding paragraphs may be used for many purposes, some of which are indicating breaks in a paper mill, automatic weighing, counting of inanimate objects on a production line, of people, and of automobiles; and operating doors.

Problems

8-1. A triode tube with characteristics as in Figs 3-15 and 3-16 is used in the circuit of Fig. 8-1. The plate-battery voltage is 100. What should be the potential between grid and cathode for plate currents of 2 and 18 ma, respectively (*a*) if the load has zero resistance, (*b*) if the load has a resistance of 3000 ohms?

8-2. The same tube is to be used in the circuit of Fig 8-2 Assume the plate current to flow in half sine-wave pulses and that the *crest values* of these pulses are determined by the curves of Figs 3-15 and 3-16. The emf between points 2 and 3 is 50 volts, rms. (*a*) What must be the rms potential between points 1 and 2 if the plate current is to be just zero when the potentiometer is at point 1? What is the average current flowing when the potentiometer is at point 2, (*b*) if the resistance of the load is zero, (*c*) if the resistance of the load is 3000 ohms?

8-3. The triode tube of Prob 8-2 is to be used in the circuit of Fig. 8-10. The resistance of the relay is 3000 ohms. The potential between points *c* and *d* is 60 volts, rms. (*a*) What must be the rms potential between points *a* and *c* to provide a direct current through the relay of 1 ma with the switch closed? (*b*) What must be the potential between points *b* and *c* to provide a direct current through the relay of 5 ma when the switch is open? (Assume no condenser across relay.)

8-4. A certain thyratron tube, which will deionize within 1000 μsec after cessation of plate current flow and has a tube drop of 12 volts under load, is used in the circuit of Fig. 8-14. If the load is a resistance of 100 ohms and the d-c source in the plate circuit has a voltage of 125, what is the minimum size of C which will ensure that the grid will assume control when S_1 is closed? (To be on the safe side assume that the tube may fire as soon as the anode is more positive than the cathode.)

8-5. (*a*) Repeat Prob. 8-4 for a d-c source voltage of 250. (*b*) Repeat Prob. 8-4 for a load resistance of 500 ohms.

8-6. A thyratron is used in the circuit of Fig. 8-15. The rms voltage between points c and d is 300. The characteristics of the tube are given in Fig. 4-25. The resistance of the load is 400 ohms, and the tube drop is 15 volts. If the condensed mercury temperature is 40°C, what is the minimum rms value of the voltage between the cathode and point a that will prevent the tube from firing?

8-7. A thyratron is used in the circuit of Fig. 8-19. The three-phase supply is at a potential of 300 volts rms, and the resistance of the load is 200 ohms. Determine the average current through the thyratron for each 30 deg of grid phase shift. Neglect tube drop and assume that the voltage supplied by the phase-shifting circuit is high enough so that firing takes place at substantially the same instant the grid voltage passes through zero.

8-8. In the circuit of Fig. 8-20 assume that the condenser has a capacitance of 7.5 μf, and the load resistance is 100 ohms. Compute the value of the adjustable resistance required to provide a grid phase shift of (*a*) 45°, (*b*) 90°. Assume that the transformer windings are so wound and connected as to give operation as in Fig. 8-21*b*, not *d*. Neglect the effect of any grid current and of the magnetizing current of the transformer.

8-9. For the circuit of Fig. 8-24 compute the output voltages for a grid phase shift of (*a*) 30° and (*b*) 60°. Obtain output voltage in per cent of maximum voltage (obtainable with grid and plate in phase). Assume that the load current is larger than the critical minimum for the choke used, and neglect the drop in the choke, tubes, and transformers.

8-10. A phototube with characteristics as in Figs. 5-2 and 5-3 is used in the circuit of Fig. 8-35. The triode has characteristics as in Figs. 3-15 and 3-16. The relay has a resistance of 5000 ohms. The battery voltage in the plate circuit of the triode is 80 volts; that in the phototube circuit is 00 volts. The grid leak R_g is 1,000,000 ohms. The relay operates on 5 ma and releases on 1 ma. (*a*) What should be the bias on the triode? (*b*) How much light must be impressed on the phototube to operate the relay?

8-11. A phototube with characteristics as in Fig. 5-5 is used in the circuit of Fig. 8-36 using a thyratron with characteristics as in Fig. 4-33. The rms voltages supplied across the divider are as follows: $cd = 6$, $dc = 60$, $bc = 300$. The total resistance in the plate circuit is 600 ohms. The resistance R_g is 500,000 ohms. How much light must be impressed on the phototube to cause the thyratron to fire?

CHAPTER 9

AUDIO-FREQUENCY AMPLIFIERS

High-vacuum tubes may be used as amplifiers, since a signal of little or no power may be applied between grid and cathode to control a comparatively large flow of power in the plate circuit. If the grid is maintained more negative than the cathode, it will not attract electrons and the grid current will be substantially zero; thus the tube will demand almost no input power. At the same time the voltage appearing across the plate output impedance is, in general, greater than that impressed on the grid, giving the tube some of the characteristics of a transformer. The ability of the tube to provide an increase or amplification of power has made it one of the outstanding contributions to the electrical art of recent years, since a transformer may increase the voltage but not the power, whereas the vacuum tube will do both.

Basically the high-vacuum tube is of the nature of a valve. Its apparent amplification of power is due to the action of its grid in controlling the flow of d-c power in the plate circuit. Thus if a signal representing a few microwatts of power is to be amplified, it is applied to the grid of a tube through a suitable circuit, such as described later in this chapter, and controls the flow of power from a d-c source which is applied to the anode of the tube through a suitable load impedance. If the signal to be amplified is alternating in nature (as is most commonly the case), the d-c plate power flow will be varied in such a manner as to produce an alternating component in the output which may, under proper circuit conditions, be of the same wave shape as the impressed signal. This control action of the vacuum tube has made possible many far-reaching developments such as talking pictures, radio communication, long-distance telephony, television, and radar.

Gas-filled tubes are not commonly used as amplifiers. The grids of most gas-filled tubes are incapable of instantaneously controlling the magnitude of the plate current, their function being limited to that of starting the flow of current at a given time.

Thus gas-filled tubes are essentially "on-off" control devices, rather than amplifiers, and are used principally in industrial applications such as were treated in Chap. 8.

Classes of Amplifiers. Amplifiers are most commonly divided into *class A*, *class AB*, *class B*, and *class C* amplifiers. Appendix A gives the definitions of these various classes, as defined by the Institute of Radio Engineers. Briefly, class A amplifiers are those in which the plate current flows throughout the grid voltage cycle and is of essentially the same wave shape as the signal impressed on the grid. The plate current in class B amplifiers flows for approximately one-half of each cycle; that of class C amplifiers flows for appreciably less than one-half of each cycle. Class AB amplifiers are intermediate between class A and class B. Occasionally the subscripts 1 and 2 are used, as class A_1, class B_2, etc., the subscript 1 indicating that the grid is negative at all times throughout its cycle of operation, and 2 indicating a positive grid swing during a short portion of each cycle.

Amplifiers may also be classed as *audio-frequency*, *radio-frequency*, *video-frequency*, or *direct-current*. Audio-frequency amplifiers are those designed to amplify frequencies within the audible spectrum, *i.e.*, approximately from 15 to 15,000 cycles, and are usually class A or, when two tubes are used in push-pull, class AB or class B. Radio-frequency amplifiers are those designed to amplify any signal of frequency higher than the audible spectrum and, except in radio receivers, are commonly class C or, when the signal is modulated, class B. The latter are often referred to as *linear amplifiers* when used in such service. Video-frequency amplifiers are those designed to amplify the frequency band required for television, *i.e.*, up to perhaps 6 Mc.

Amplifiers may also be classified as *power amplifiers* or *voltage amplifiers*. Although all amplifiers are, in the strict sense of the word, power amplifiers (else they would not be *amplifiers*), voltage amplifiers are used in circuits where only a small part of the maximum power output of the tube is required but a considerable voltage gain is desired. Transformers may not be capable of supplying the need, either because some power gain is necessary or because the frequency response or input impedance requirements may be such as to make the design of a suitable transformer impractical. Vacuum tubes have a much higher input impedance than is obtainable with any reasonable design of transformer and

thus are ideal for increasing the voltage from a high-impedance source, such as so often is encountered in communication work. With suitable circuits they may also be used to amplify a current or to cause a given voltage to produce an amplified current.

Distortion. In most applications of a-f amplifiers it is desired that the changes in plate current be exactly proportional to the changes in applied grid potential; and this should be true, within certain prescribed limits, no matter how rapidly these changes occur or how great may be their magnitude. If two or more sinusoidal emfs of different frequencies are simultaneously impressed on the grid of a triode amplifier, as in Fig. 9-1, the induced voltage across the load should contain exactly the same frequency components in the same relative phase positions, each amplified to the same degree; also, if the impressed emfs are varied in magnitude, the components appearing in the output should vary in magnitude in a direct ratio. Unless this is true, distortion is present.

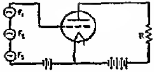


FIG. 9-1 More than one frequency component is generally applied to the grid of a triode amplifier at any given time

Distortion may be divided into three classifications:

1. Amplitude or harmonic distortion.
2. Frequency distortion
3. Phase distortion.

Amplitude distortion is said to exist when the ratio of output to input voltage varies as the amplitude of the impressed signal varies. It is characterized by the introduction of new components (or harmonics) having frequencies equal to multiples of the impressed frequencies and to the sums and differences of these frequencies and their harmonics. Amplitude distortion is caused by any nonlinear circuit element, usually the tube since the characteristic curves of vacuum tubes are not straight. It may be kept small by operating the tube on the straightest portion of the characteristic curve and by the use of external circuit elements of such magnitude as to minimize the effect of tube curvature.

Frequency distortion is said to exist when the ratio of output to input voltage varies as the frequency of the impressed signal varies, its amplitude remaining constant. No new frequency components are introduced, but the relative magnitudes of the

various components are changed. Frequency distortion is caused by circuit impedances which vary in magnitude with frequency. These impedances are largely in the external circuit associated with the tube, but the capacitances between the electrodes of the tube also have an appreciable effect at the higher frequencies.

Phase distortion is said to exist when the relative phases of the various frequency components are different in the output than they are in the input. This phase shift between the various components is due to a difference in time delay through the circuit as the frequency is varied, the phase shift being zero if all components are delayed by the same amount. That this is true may be seen by noting that equal time delay will cause all components of a given portion of a complex impressed wave to arrive at the output in exactly the same relation, one to the other, as they occupied at the input. At low frequencies phase distortion is due largely to reactance in the external circuit elements of the amplifier, but at very high radio frequencies the transit time of the electron in passing through the tube also causes appreciable time delay. Phase distortion may be kept small by designing the circuit of the amplifier to keep the delay of all components approximately equal. This frequently means that the delay is kept as near zero as possible.

Decibels. Amplifier gain is commonly measured in *decibels*, abbreviated db. The decibel is essentially a power *ratio* unit as indicated by the defining equation[1]

$$\text{db} = 10 \log_{10} \frac{P_2}{P_1} \tag{9-1}^*$$

where P_2 and P_1 are the output and input power, respectively. Since $P = E^2/R$, Eq. (9-1) may be written

$$\text{db} = 10 \log_{10} \frac{E_2^2}{E_1^2} \frac{R_1}{R_2}$$

or, *if the resistances are equal*,

$$\text{db} = 20 \log_{10} \frac{E_2}{E_1} \tag{9-2}^*$$

[1] The asterisk (*) after an equation indicates that the equation is a final result rather than a preliminary equation leading to a mathematical conclusion. At subsequent points in this chapter equations marked with an asterisk will follow a number of preliminary equations not so marked. (See preface to the first edition.)

The decibel is often used as a unit of actual power output by giving the number of decibels above a given reference level. The reference level has varied with the application, resulting in considerable confusion, but the level most commonly used in broadcast and sound work has been 1 mw. Thus an output of −20 db indicates an output of 20 db below 1 mw. If these values are inserted in Eq. (9-1), the result will be

$$-20 = 10 \log_{10} \frac{P_2}{0.001}$$

which, when solved for P_2, shows that the power output is 0.00001 watt, or 10 μw.

Input Admittance of a Vacuum Tube. It has been stated that the impedance between grid and cathode of a tube, when used as an amplifier, should be as high as possible. The grid is normally maintained negative to prevent the flow of electrons to that electrode and so keep the impedance high. However, even with a negative grid it is probable that neither the resistive nor the reactive components will be even approximately infinite owing to the capacitances between the grid and the other electrodes in a tube. These capacitances are usually small, being of the order of 2 to 8 $\mu\mu$f each, but the grid-plate capacitance is in series with the plate voltage, which materially increases its effect on the input circuit, as will be shown.

The circuit of a simple amplifier using a triode tube is shown in Fig. 9-2*a* and the complete equivalent a-c circuit, including the capacitances, in (*b*), assuming the grid to be maintained negative throughout the cycle. Since we wish to know the admittance as seen from the generator $\mathbf{E}_g$ we must first determine the current $\mathbf{I}_g$, after which the admittance may be found from $\mathbf{Y} = \mathbf{I}_g/\mathbf{E}_g$.

The potential difference between points *A* and *B*, Fig. 9-2*b*, is $\mathbf{E}_p$, the plate voltage. In succeeding sections of this chapter we shall develop equations for this potential in terms of $\mathbf{E}_g$ for the various types of amplifiers; therefore, it will simplify our present problem to replace that entire portion of the circuit of Fig. 9-2*b* which is contained in the dashed lines, with a single generator having a voltage $\mathbf{E}_p$. Since $\mathbf{E}_p$ can be expressed in terms of $\mathbf{E}_g$ we may write $\mathbf{E}_p = \mathbf{A}\mathbf{E}_g$ where $\mathbf{A}$ is a vector having a magnitude A equal to the ratio of the magnitudes of the plate and grid voltages and a phase angle α equal to the difference in phase between the

two voltages; therefore, we may write $\mathbf{A} = A\angle\alpha$. Referring to Fig. 3-26 it will be seen that α will have a value of 180 deg. for a resistance load. It seems preferable to use the angle ψ between $-\mu\mathbf{E}_g$ and $\mathbf{E}_p$ which is zero for a resistance load and is equal to $(\theta - \phi)$ for reactive loads (see Figs. 3-28 and 3-30). We may,

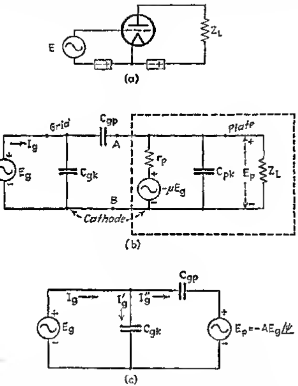


FIG. 9-2. (a) Triode amplifier with load impedance Z. (b) and (c) Equivalent circuits of (a) for determining the input admittance.

therefore, write $\mathbf{A} = A\angle 180° + \psi = -A\angle\psi$. The circuit of Fig. 9-2b may now be replaced by that of (c).

The input admittance of the tube may now be determined by solving for the current $\mathbf{I}_g$ in terms of the circuit constants and dividing this result by $\mathbf{E}_g$. We may write[1]

[1] Throughout the rest of this book vectors and complex quantities are indicated by the use of boldface type.

$$\mathbf{I}_g = \mathbf{I}'_g + \mathbf{I}''_g \tag{9-3}$$

$$\mathbf{I}'_g = j\omega C_{gk}\mathbf{E}_g \tag{9-4}$$

$$\mathbf{I}''_g = (\mathbf{E}_g - \mathbf{E}_p)j\omega C_{gp} \tag{9-5}$$

or

$$\mathbf{I}''_g = (\mathbf{E}_g + A\mathbf{E}_g\underline{/\psi})j\omega C_{gp} \tag{9-5a}$$

Equation (9-5*a*) may be handled most readily by changing into rectangular coordinates,

$$\mathbf{I}''_g = j\omega C_{gp}\,[\mathbf{E}_g + A\mathbf{E}_g(\cos\psi + j\sin\psi)] \tag{9-6}$$

Substituting Eqs. (9-6) and (9-4) into Eq. (9-3) and dividing by $\mathbf{E}_g$ gives

$$\mathbf{Y}_g = \frac{\mathbf{I}_g}{\mathbf{E}_g} = -A\omega C_{gp}\sin\psi + j\omega(C_{gk} + C_{gp} + AC_{gp}\cos\psi) \tag{9-7}^*$$

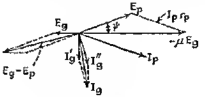
Inspection of Eq. (9-7) shows that both the in-phase and quadrature components of the input admittance are affected by the plate circuit. Furthermore, if $\sin\psi$ is positive, the in-phase component is negative. This means that energy is being fed back through the grid-plate capacitance in the proper phase to supply, at least partially, the losses in the circuit. With suitable input and output circuits an alternating emf may be set up between grid and cathode and sustained by this feed-back action, producing a useful output in the plate circuit (see discussion on the tuned-grid-tuned-plate oscillator in Chap. 11, page 445).

A vector diagram of the currents and potentials of the tube is shown in Fig. 9-3, (*a*) being for a lagging plate current and (*b*) for a leading current. The component $\mathbf{I}'_g$ of the grid current is shown leading the grid voltage by 90 deg according to Eq. (9-4) while $\mathbf{I}''_g$ leads the voltage $(\mathbf{E}_g - \mathbf{E}_p)$ by 90 deg according to Eq. (9-5). The resultant grid current $\mathbf{I}_g$ is seen to have an in-phase component which is negative with respect to the grid voltage in (*a*), indicating feedback of energy as stated in the preceding paragraph, while the in-phase component of grid current is positive in (*b*) indicating consumption of power by the grid circuit.

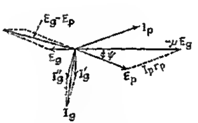
The term in the parentheses of Eq. (9-7) is the *equivalent input capacitance* C_g of the triode or

$$C_g = C_{gk} + (1 + A\cos\psi)\,C_{gp} \tag{9-8}^*$$

Equation (9-8) shows that the effect of the grid-plate capacitance C_{gp} is increased by the factor $(1 + A)$ under the common condition of the plate voltage being in phase with $-\mu \mathbf{E}_g$. Since A is usually many times greater than unity, the equivalent input capacitance of a triode is much larger than its geometric capacitance $(C_{gk} + C_{gp})$. For example, suppose a triode to have the



(a) Lagging power factor



(b) Leading power factor

FIG. 9-3. Vector diagrams of the circuit of Fig. 9-2c.

following interelectrode capacitances: $C_{gk} = 4\mu\mu f$, $C_{gp} = 6\ \mu\mu f$; and suppose the gain of the tube, with a given resistance load, to be 5. The equivalent input capacitance is then

$$C_g = 4 + (1 + 5)6 = 40\ \mu\mu f$$

which is considerably more than the geometric capacitance of 10 $\mu\mu f$. The reactance of a 40-$\mu\mu f$ condenser at a frequency of only 10,000 cycles/sec is but 400,000 ohms, and at radio frequencies it certainly cannot be neglected.

A more general form of Eq. (9-8), applicable to multigrid tubes as well as to triodes, is

$$C_e = C_{in} + AC_{gp} \cos \psi \tag{9-9}^*$$

Here C_{in} is the geometric input capacitance of the tube; *i.e.*, it is the total capacitance between the grid and all other electrodes of the tube, including the cathode, the other electrodes being tied together. In the triode $C_{in} = C_{gk} + C_{gp}$ since the cathode and plate are the only other electrodes in the tube. In a pentode this capacitance must include the effect of the screen and suppressor grids as well. Tube manuals give C_{in} for most tubes.

In pentodes the capacitance C_{gp} is very small owing to the presence of the screen and suppressor grids so that the equivalent input capacitance C_e is very nearly equal to the geometric capacitance C_{in}. As an example, consider a typical pentode in which $C_{in} = 8$ $\mu\mu$f and $C_{gp} = 0.005$ $\mu\mu$f. If the gain of the tube with a given resistance load is 150 (a typical value for pentode amplifiers) Eq. (9-9) gives

$$C_e = 8 + 150 \times 0.005 = 8.75 \ \mu\mu\text{f}$$

which is only slightly higher than the geometric capacitance of 8 $\mu\mu$f. Incidentally, if a gain of 150 were possible with the triode tube of the preceding example, instead of 5 as assumed, C_e for the triode would have been 910 $\mu\mu$f.

1. VOLTAGE AMPLIFIERS

A voltage amplifier was defined in an earlier section of this chapter as one designed to produce a large increase in voltage, with power output considerations being secondary. Voltage amplifiers may be classified by the type of coupling used between tubes as *resistance-coupled*, *transformer-coupled*, and *impedance-coupled* amplifiers. Those described in this chapter are all class A and are designed to amplify audio and video frequencies only.

Resistance-coupled Amplifiers. The distinguishing feature of a resistance-coupled amplifier is that a resistance is inserted in the plate lead of the amplifier tube and the output voltage is developed across this resistance. The basic circuit of a resistance-

coupled amplifier is shown in Fig. 9-4, in which tube 1 is the amplifier tube under consideration and R_1 is the resistance in series with the plate of the tube, across which the output voltage is developed. While batteries are not normally used in commercial amplifiers, they are shown here to simplify the circuit for the purposes of discussion. Practical circuits in which the direct voltages for both plate and grid are supplied by a single rectifier are shown later (Fig. 9-8).

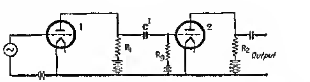


FIG. 9-4. Basic circuit of a two-stage, resistance-coupled amplifier.

The output voltage is applied to the grid of tube 2 through the condenser C.[1] This condenser prevents the direct voltage in the plate circuit of tube 1 from affecting the grid of tube 2 but permits the alternating component to be impressed on the grid. As the input impedance of tube 2 is very high, this condenser need not be large; a reactance of some tens of thousands of ohms will be negligible as compared with the input impedance of tube 2 and the paralleling resistance R_g. The resistance R_g, generally known as the *grid leak*, provides a path from the grid of the second tube to the C bias to maintain the grid negative and drain off any electrons collecting on the grid during a momentary positive swing under large excitation voltages. It is usually of the order of 0.5 megohm.

Equivalent Circuits of Resistance-coupled Amplifiers. The performance of resistance-coupled amplifiers may be analyzed best by drawing the equivalent circuit (Fig. 9-5a). The tube is replaced by a source of alternating voltage $-\mu E_g$ and a resistance r_p (see pages 60 to 64 for proof of the legitimacy of this substitution). Capacitances C_1 and C_2 are intended to represent the internal capacitances C_p of the plate circuit of tube 1 and C_g of the grid circuit of tube 2, respectively, together with any capacitance

[1] Tube 2 is also shown as resistance-coupled, although any other type of coupling may be employed.

external to the tube, as in the wiring. In triodes the capacitance C_2 is much larger than the geometric electrode capacitance C_{in}, owing to reaction from the plate circuit [see Eq. (9-9)], but C_2 is usually equal to C_{out} in all tubes.

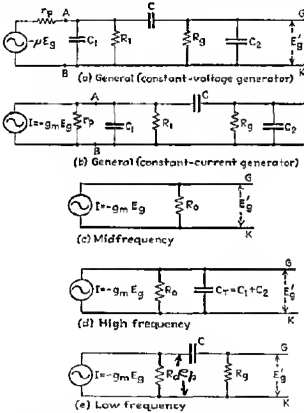


FIG. 9-5 Equivalent circuits of a resistance-coupled amplifier.

The solution of this circuit can be somewhat simplified by changing the constant-voltage generator of Fig. 9-5a to an equivalent constant-current generator. This may be done by applying Norton's theorem to that portion of the circuit lying to the left of points

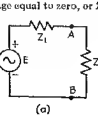
AB.[1] This will give a constant-current generator delivering a current $\mathbf{I} = -\mu \mathbf{E}_g/r_p = -g_m \mathbf{E}_g$, with an internal shunting impedance of r_p. With this change the circuit of Fig. 9-5a becomes that of Fig. 9-5b.

Most resistance-coupled amplifiers are designed to amplify components of all frequencies equally throughout a given frequency spectrum. Thus it is necessary that the capacitances and resistances constituting the output circuit of the tube be of such magnitude that the capacitances (including those of the associated tubes) have negligible effect throughout the desired range of constant gain. Within this limited frequency range, therefore, the circuit may be simplified by eliminating the capacitances C_1 and C_2 and replacing C with a short circuit, whence the constant-current generator works into a resistance R_0 equal to the three resistances r_p, R_1, and R_g in parallel (Fig. 9-5c) where

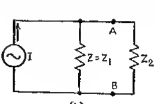
$$R_0 = \frac{1}{(1/r_p) + (1/R_1) + (1/R_g)} = \frac{r_p R_1 R_g}{r_p R_1 + r_p R_g + R_1 R_g} \qquad (9\text{-}10)$$

[1] Norton's theorem states that any network of generators and impedances supplying a given load impedance may be replaced by a simple generator delivering a current $\mathbf{I}$ and having an internal shunting impedance $\mathbf{Z}$, where $\mathbf{I}$ is the current flowing when the terminals of the network are short-circuited and $\mathbf{Z}$ is the impedance measured looking back into the network with all the generators replaced by impedances equal to their internal impedances. (For proof of this theorem see books such as W. L. Everitt, "Communication Engineering," McGraw-Hill Book Company, Inc., New York, 1937.)

As an example of the application of this theorem consider the circuit of Fig. a of this footnote. Norton's theorem may be applied by considering $\mathbf{Z}_2$ as the load and the rest of the circuit as the network of generators and impedances. Application of Norton's theorem will give the circuit of Fig. b in which the generator delivers a current $\mathbf{I}$ equal to the current flowing when the terminals A and B are shorted in the original circuit of Fig. a, or $\mathbf{I} = \mathbf{E}/\mathbf{Z}_1$, and the shunting impedance $\mathbf{Z}$ in Fig. b is that which is seen looking toward the left from terminals AB in Fig. a with the generator voltage equal to zero, or $\mathbf{Z} = \mathbf{Z}_1$.



(a)



(b)

At frequencies above the desired range, the effect of the capacitances C_1 and C_2 is appreciable, whereas that of C continues to be negligible, and the circuit reduces to that of Fig. 9-5*d*; at frequencies below the desired range, C_1 and C_2 have negligible effect, but C must be considered, giving the circuit of Fig. 9-5*e* where

$$R_e = \frac{1}{(1/r_p) + (1/R_1)} = \frac{r_p R_1}{r_p + R_1} \tag{9-11}$$

Computations of amplifier gain are simplified by the use of these circuits, since a few simple calculations will show which, if any, of the capacitances are negligible at a given frequency and, therefore, which simplified circuit of Fig. 9-5 should be used.

The curve of Fig. 9-6 shows the variation in gain with frequency of a resistance-coupled amplifier using a triode tube. The effect of the capacitance C at the lower frequencies and of the capacitances C_1 and C_2 at the higher frequencies may be clearly seen. The curve is nearly flat over a very wide frequency range, a principal advantage of this type of coupling.

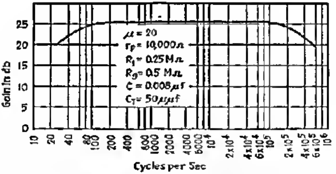


FIG. 9-6 Gain vs frequency characteristic of a resistance-coupled amplifier.

Gain of a Resistance-coupled Amplifier. The gain of voltage amplifiers may be expressed in terms of decibels as

$$\text{Gain in db} = 20 \log_{10} \frac{E_g'}{E_g} \tag{9-12}$$

where E_g' and E_g are the magnitudes of the voltages impressed on the grids of tubes 2 and 1 of Fig. 9-4, respectively, or it may

be expressed as the simple ratio of the two grid voltages

$$A = \frac{E_g'}{E_g} \tag{9-13}$$

where A is the voltage gain of the amplifier expressed as a ratio.

Equation (9-13) gives no indication of the phase difference between $\mathbf{E}_g'$ and $\mathbf{E}_g$. In the circuit of Fig. 9-4 the output and input circuits are entirely independent of each other, and the relative phase of the two voltages is unimportant. Information regarding the phase shift is of great importance, however, in feed-back amplifiers and in amplifiers used for television purposes. Therefore, it seems desirable to express the amplifier gain as a complex number,

$$\mathbf{A} = \frac{\mathbf{E}_g'}{\mathbf{E}_g} \tag{9-13}$$

where the voltages are vectors as indicated by boldface type.

The gain of a resistance-coupled amplifier may be determined by direct solution of the networks of Fig. 9-5a or b, but such a procedure is laborious and the work may be more readily carried out with the aid of the simplified circuits of c, d, and e, Fig. 9-5. In Fig. 9-5c the output voltage $\mathbf{E}_g'$ is equal to the current times the resistance R_0 or

$$\mathbf{E}_g' = \mathbf{I}R_0 = (-g_m\mathbf{E}_g)R_0 \tag{9-15}$$

Solving for $\mathbf{A}$ gives

$$\mathbf{A}_m = \mathbf{E}_g'/\mathbf{E}_g = -g_mR_0 = g_mR_0\underline{/180°} \tag{9-16}^*$$

where the subscript m under $\mathbf{A}$ indicates that this equation applies only to the mid-frequency band to which the circuit of Fig. 9-5c is applicable.

Equation (9-16) shows that there is no phase shift throughout the frequency range to which the circuit of Fig. 9-5c is applicable, except for the 180 deg phase reversal which takes place in the tube. It also shows that the gain throughout this range is independent of frequency, corresponding approximately to the frequency range 150 to 100,000 cycles/sec in Fig. 9-6.

The gain at frequencies above the flat mid-frequency range of Fig. 9-6 may be found from the circuit of Fig. 9-5d. From this figure we may write

$$\mathbf{E}_g' = \mathbf{I}\frac{1}{(1/R_0) - (1/jX_T)} = -g_m\mathbf{E}_g\frac{R_0X_T}{X_T + jR_0} \tag{9-17}$$

where $X_T = 1/\omega C_T$. The gain is therefore

$$\mathbf{A}_h = \frac{\mathbf{E}'_g}{\mathbf{E}_g} = -g_m R_0 \frac{X_T}{X_T + jR_0} \tag{9-18}$$

or

$$\mathbf{A}_h = g_m R_0 \frac{X_T/180°}{\sqrt{X_T^2 + R_0^2}/\tan^{-1}(R_0/X_T)} \tag{9-19}$$

$$\mathbf{A}_h = g_m R_0 \frac{1}{\sqrt{1 + (R_0^2/X_T^2)}} \Big/ 180° - \tan^{-1}\frac{R_0}{X_T} \tag{9-20}*$$

where the subscript h under $\mathbf{A}$ indicates that these equations are applicable only throughout the h-f[1] range to which the circuit of Fig. 9-5d applies. Inspection of Eq. (9-20) shows that the magnitude term, when plotted against frequency, will produce a curve similar to that of the h-f end of the curve of Fig. 9-6. The phase shift in the amplifier is seen to lag behind the 180-deg phase reversal taking place in the tube, by an amount that increases with frequency, approaching 90 deg as a limit.

For frequencies below the flat mid-frequency region of Fig. 9-6, the circuit of Fig. 9-5e must be used. For the voltage across R_e we may write

$$\mathbf{E}_r = \mathbf{I}\frac{1}{\dfrac{1}{R_e} + \dfrac{1}{R_g - jX_c}} = -g_m\mathbf{E}_g \frac{R_e(R_g - jX_c)}{R_e + R_g - jX_c} \tag{9-21}$$

where $X_c = 1/\omega C$. The current through R_g is

$$\mathbf{I}_g = \frac{\mathbf{E}_r}{R_g - jX_c} \tag{9-22}$$

The output voltage $\mathbf{E}'_g$ is evidently

$$\mathbf{E}'_g = \mathbf{I}_g R_g \tag{9-23}$$

Substituting Eq. (9-21) into (9-22) and substituting the resulting expression into Eq. (9-23) give

$$\mathbf{E}'_g = -g_m\mathbf{E}_g \frac{R_e R_g}{R_e + R_g - jX_c} \tag{9-24}$$

[1] Throughout this chapter the abbreviations "h-f" and "l-f" have been used for "high-frequency" and "low-frequency," respectively.

To simplify, let

$$R_a + R_g = R_e \tag{9-25}$$

Also, since R_0 is the resistance of r_p, R_1, and R_g in parallel and R_a is the resistance of r_p and R_1 in parallel, it is evident that R_0 is the resistance of R_a and R_g in parallel or

$$R_0 = \frac{R_a R_g}{R_a + R_g} = \frac{R_a R_g}{R_e} \tag{9-26}$$

Solving Eq. (9-26) for $R_a R_g$ and substituting this value and Eq. (9-25) into Eq. (9-24) give

$$\mathbf{E}_g' = -g_m \mathbf{E}_g R_0 \frac{R_e}{R_e - jX_c} \tag{9-27}$$

We may now write the equation for gain, in polar form, as

$$\mathbf{A}_l = \frac{\mathbf{E}_g'}{\mathbf{E}_g} = g_m R_0 \frac{R_e \underline{/180^\circ}}{\sqrt{R_e^2 + X_c^2} \underline{/-\tan^{-1}(X_c/R_e)}} \tag{9-28}$$

or

$$\mathbf{A}_l = g_m R_0 \frac{1}{\sqrt{1 + (X_c^2/R_e^2)}} \underline{\Big/ 180^\circ + \tan^{-1}\frac{X_c}{R_e}} \tag{9-29}^*$$

The subscript l under $\mathbf{A}$ indicates that this equation applies only to the l-f range to which the circuit of Fig. 9-5c applies. Inspection of Eq. (9-29) shows that the magnitude term, when plotted against frequency, will give a curve similar to the l-f end of the curve of Fig. 9-6. The angle in Eq. (9-29) shows that at the lower frequencies the phase shift exceeds the normal 180 deg produced in the tube by an angle that approaches 90 deg as the frequency approaches zero.

A study of Eqs. (9-16), (9-20), and (9-29) discloses that all contain the common term $g_m R_0$. It may be seen that the term under the radical in Eq. (9-20) will approach unity as the frequency drops toward the middle or desired range, whereas that of Eq. (9-29) approaches unity as the frequency rises. It is, therefore, possible to write a single equation for the gain of a resistance-coupled amplifier that applies to all frequencies by combining

Eqs (9-16), (9-20) and (9-29) to give

$$\mathbf{A} = \underbrace{g_m R_0}_{\substack{\text{Mid}\\ \text{frequency}\\ \text{gain}}} \underbrace{\frac{1}{\sqrt{1 + (R_0^2/X_T^2)}}}_{\substack{\text{High frequency}\\ \text{multiplying}\\ \text{factor}}} \underbrace{\frac{1}{\sqrt{1 + (X_c^2/R_e^2)}}}_{\substack{\text{Low frequency}\\ \text{multiplying}\\ \text{factor}}}$$

$$\Big/ \underbrace{180^\circ}_{\substack{\text{Tube phase}\\ \text{reversal}}} - \underbrace{\tan^{-1}\frac{R_0}{X_T}}_{\substack{\text{High frequency}\\ \text{phase shift}}} + \underbrace{\tan^{-1}\frac{X_c}{R_e}}_{\substack{\text{Low frequency}\\ \text{phase shift}}} \qquad (9\text{-}30)$$

where $\mathbf{A}$ = vector voltage gain of single-stage amplifier
g_m = mutual conductance of driving tube
X_T = reactance of condenser C_T
X_c = reactance of coupling condenser C
r_p = plate resistance of driving tube
R_0 = value given by Eq (9-10)
R_e = value given by Eq. (9-25)

Frequency Response of Resistance-coupled Amplifiers. There are two requirements to be met by most resistance-coupled amplifiers. (1) They must produce uniform response over a given frequency band (see Fig. 9-6) and (2) they should produce the maximum possible gain consistent with securing the desired frequency response. Equation (9-30) shows that the maximum gain is $g_m R_0$ which is the gain in the mid-frequency range. If the amplifier is to have a flat response curve, it is obvious that the gain throughout the entire desired frequency band should be equal to $g_m R_0$. This will be the case if R_0/X_T and X_c/R_e are small compared with unity. Therefore, $R_0/X_T \ll 1$ at the highest frequency that it is desired to amplify, and $X_c/R_e \ll 1$ at the lowest frequency. The upper and lower frequency limits may evidently be extended by *decreasing C_T and increasing C, respectively*, but there are obvious limits to this course of action, especially at the *upper frequency limit*. It is possible to achieve the same results by decreasing R_0 and increasing R_e, which may be done for a given tube (therefore, for a given value of r_p) by decreasing R_1 and increasing R_g. The decrease in R_0 evidently results in a loss in gain, but this is the price that must often be paid for a good frequency response.

Equations may be derived for the maximum and minimum

frequencies above and below which the gain is less than a desired fraction of the mid-frequency gain. If this fraction is K, we may write $A_l/A_m = K$ at the minimum frequency and $A_h/A_m = K$ at the maximum frequency. After substituting the values of A_m, A_h, and A_l from the magnitude terms of Eqs. (9-16), (9-20), and (9-29) into these ratios we may solve for these minimum and maximum frequencies, finding for the minimum frequency

$$f_{min} = \frac{K}{2\pi R_e C\sqrt{1 - K^2}} \qquad (9\text{-}31)^*$$

and for the maximum frequency

$$f_{max} = \frac{\sqrt{1 - K^2}}{2\pi K R_0 C_T} \qquad (9\text{-}32)^*$$

A common value for K is 0.707 (3 db down from the mid-frequency gain) which means that the output voltage at the minimum and maximum frequencies has dropped to 70.7 per cent of the mid-frequency value. Substituting 0.707 for K in Eqs. (9-31) and (9-32) gives for the minimum frequency for a 3-db drop in gain

$$f_1 = \frac{1}{2\pi R_e C} \qquad (9\text{-}33)^*$$

and for the maximum frequency for a 3-db drop in gain

$$f_2 = \frac{1}{2\pi R_0 C_T} \qquad (9\text{-}34)^*$$

The frequencies f_1 and f_2 are often known as the half-power points because the power output at these frequencies is half the power output in the mid-frequency range. When the half-power points have been determined from the two preceding equations, it may be said that the amplifier gain is flat between these two limits.

Equation (9-34) may be rewritten as

$$R_0 = \frac{1}{2\pi f_2 C_T}$$

which shows that the upper half-power point occurs at a frequency for which the reactance of the shunting capacitance equals the equivalent resistance R_0. A similar analysis of Eq. (9-33) shows that the reactance of the series-blocking condenser C is equal to the resistance R_e at the lower half-power point.

Universal Amplification Curves. Equation (9-30) may be rewritten by replacing R_s and R_0 with expressions obtained from Eqs. (9-33) and (9-34) and expressing X_T and X_c in terms of C_T and C, respectively, as follows: $R_0 = 1/2\pi f_2 C_T$, $R_s = 1/2\pi f_1 C$, $X_T = 1/2\pi f C_T$, and $X_c = 1/2\pi f C$, where f_1 and f_2 are as defined in the preceding section and f is the frequency at which the performance of the amplifier is to be determined. When these substitutions are made, Eq. (9-30) reduces to

$$A = \underbrace{g_m R_0}_{\substack{\text{Mid}\\ \text{frequency}\\ \text{gain}}} \underbrace{\frac{1}{\sqrt{1 + (f/f_2)^2}}}_{\substack{\text{High frequency}\\ \text{multiplying}\\ \text{factor}}} \underbrace{\frac{1}{\sqrt{1 + (f_1/f)^2}}}_{\substack{\text{Low frequency}\\ \text{multiplying}\\ \text{factor}}} \Bigg/ \underbrace{180°}_{\substack{\text{Tube phase}\\ \text{reversal}}} - \underbrace{\tan^{-1}\frac{f}{f_2}}_{\substack{\text{High-frequency}\\ \text{phase shift}}} + \underbrace{\tan^{-1}\frac{f_1}{f}}_{\substack{\text{Low frequency}\\ \text{phase shift}}} \quad (9\text{-}35)^*$$

The l-f and h-f multiplying factors of Eq. (9-35) are plotted in Fig. 9-7a against f_1/f or f/f_2 as the abscissa, and the phase shift

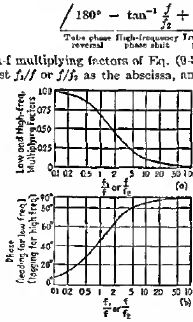


FIG. 9-7. Universal amplification and phase-shift curves.

over and above the normal 180-deg phase reversal of the tube is plotted in Fig. 9-7b. Application of these curves merely requires the computation of the frequencies f_1 from Eq. (9-33) and f_2

from Eq. (9-34). Numerical values may then be read from the curve to obtain the gain and phase shift of the amplifier at any desired frequency in terms of its mid-frequency performance.

Characteristics of Tubes Suitable for Use in Resistance-coupled Amplifiers. Equation (9-30) shows that maximum gain will be realized by using a tube with high g_m. A high r_p is also desirable (which means that μ should also be high, since $\mu = g_m r_p$) or at least r_p should be large compared to R_1, since this will ensure a maximum value of R_0 for a given value of R_1. Most commercial tubes suitable for a-f voltage amplifiers, both triodes and pentodes, have approximately the same g_m, being of the order of 1000 to 2000 μmhos, but pentodes have a much higher r_p than triodes and are therefore preferred for resistance-coupled amplifiers. Another reason for preferring the pentode is that its equivalent input capacitance is lower than that of the triode, since the screen and suppressor grids virtually eliminate the grid-plate capacitance (see pages 260 to 264).

Video-frequency amplifiers, such as are used in television transmitters and receivers and in radar units, must amplify such a wide band of frequencies that R_0 must be made very small, only a few thousand ohms. This is accomplished by reducing the coupling resistance R_1 to a few thousand ohms, whence $R_0 = R_1$ for all practical purposes. The gain of such amplifiers using conventional tubes is necessarily very low, but special pentode tubes have been developed for such service which have a g_m as high as 10,000 μmhos. With these tubes and with special compensating networks to improve the h-f and l-f response so that R_1 may be at least several thousands of ohms, gains of the order of 50 (34 db) per stage are obtainable over a frequency range from a few cycles per second up to several megacycles per second.

Tubes with high values of g_m are not used in ordinary resistance-coupled amplifiers designed for a-f service, even though Eq. (9-30) indicates that the gain should increase directly with g_m. A reason for avoiding them is that they draw a relatively large plate current, one such tube having a rated plate current of 10 ma. This magnitude of current flowing through a coupling resistance of 100,000 ohms, as is commonly used in a-f amplifiers, would require a plate supply voltage of $0.01 \times 100{,}000 = 1000$ plus the normal plate voltage of the tube. In a video amplifier,

on the other hand, where R_1 is perhaps 4000 ohms, the supply voltage need be only $0.01 \times 4000 = 40$ plus the tube plate voltage; therefore, the high plate current is less objectionable.

It should be evident from Eqs. (9-16) and (9-34) that, when the band width[1] of an amplifier using a given tube is increased by reducing R_1 (and, therefore, R_0), the improvement is always obtained at the expense of the mid-frequency gain. The extent of the reduction in gain is to a large extent a function of the tube used, since a tube with a higher g_m or with lower input or output capacitances is capable of producing a higher gain for a given band width.

A good figure of merit for a given amplifier is obtained by taking the product of the mid-frequency gain and the band width of the amplifier. The gain is given by Eq. (9-16) and the band width by Eq. (9-34). The product of these two expressions, disregarding the phase angle, is

$$\text{(Gain)(band width)} = \frac{g_m}{2\pi C_T} \qquad (9\text{-}36)^*$$

The capacitance C_T includes the wiring capacitances as well as the tube capacitances. The figure of merit of the tube alone is

$$\text{Figure of merit of tube} = \frac{g_m}{2\pi(C_g + C_p)} \qquad (9\text{-}37)^*$$

where C_p is the total capacitance between plate and cathode of the tube and C_g is defined by Eq. (9-9). This shows that, for a given band width, the amplifier gain may be increased by using a tube having either a higher g_m or lower electrode capacitances, or both.

Magnitudes of Circuit Elements in Resistance-coupled Amplifiers. The resistance R_1 should not be made too large, even when a very wide band width is not required, since the d-c drop through this resistance will reduce the plate voltage on the tube and therefore decrease g_m. This may be remedied by an increase in the voltage of the d-c supply, but there are practical limits to this method of solution. The resistance R_1 is usually of the order of

[1] Strictly speaking, band width is equal to $f_2 - f_1$, but f_1 is normally so small compared to f_2 that the difference between the two frequencies is practically equal to f_2. Thus it may be said that the band width is given by Eq. (9-34).

50,000 to 500,000 ohms in a-f amplifiers and a few thousand ohms in video amplifiers.

The grid-leak resistance R_g must not be too large either. A small amount of gas is present in all tubes; and if an excessive potential is suddenly applied to the grid, gas may be ionized whence the positive ions formed will constitute a reverse grid current tending to bias the tube positively. The resistance R_g is generally made as high as 250,000 to 500,000 ohms in voltage amplifiers but should be much lower for power amplifiers where ionization is more likely to occur (see later sections of this chapter).

The coupling condenser must be of good quality with very low leakage. Any leakage present will permit direct current to flow through the grid leak R_g and so tend to bias the grid of the driven tube positively. Good paper condensers are generally satisfactory unless it is desired to extend the amplification to frequencies of only a few cycles per second where the required capacitance of the condenser is so large that mica insulation must be used to keep down the leakage. The coupling condenser is usually of the order of 0.004 to 0.01 μf.

Amplifiers must also give good response to transients. If the coupling condenser and grid leak are too large, the time constant may be so great as to cause trouble. A sudden peak of voltage applied to the amplifier may cause the grid of the second tube to go momentarily positive, thus charging the coupling condenser negatively on the side toward the grid. This charge may be sufficient to send the grid voltage of this tube below cutoff as soon as the transient has passed, and the amplifier is then inoperative until the charge has leaked off. Since the time required to discharge the condenser is proportional to the product R_gC and the l-f response is a complex, inverse function of the ratio $1/R_gC$ [more particularly of the ratio X_c/R_e, see Eq. (9-29)], the requirements of good l-f response and of satisfactory transient response are incompatible, and compromises must be made.

Example. Let us assume that an amplifier is to be designed using pentode tubes with the following coefficients: $g_m = 1500$ μmhos, $r_p = 1.5$ megohms, $C_{in} = 4$ $\mu\mu$f, $C_{out} = C_p = 7$ $\mu\mu$f, $C_{gp} = 0.005$ $\mu\mu$f. C_{in} is the total capacitance measured between the control grid and the cathode of tube 2 (Fig. 9-4) and C_{out} is the total capacitance between plate and cathode of tube 1, all other electrodes being tied to the cathode during the measurements. These capacitances are usually given in tube handbooks. Let us assume that the total wiring capacitance of the circuit, in parallel with C_{in} and

C_{out}, is 10 $\mu\mu f$. Then C_T of Fig. 9-5d is $C_T = 4 + 7 + 10 = 21$ $\mu\mu f$ [assuming that the second term of Eq. (9-9) is negligibly small]. Let us assume the remaining circuit constants to be $R_1 = 100{,}000$ ohms, $R_g = 500{,}000$ ohms, $C = 0.01$ μf.

Substituting in Eq. (9-10) gives $R_0 = 79{,}000$ ohms. Next, substituting in Eq. (9-16) gives the mid-frequency gain as $A_m = 1500 \times 10^{-6} \times 79{,}000 = 118.5$. The minimum and maximum frequencies which the amplifier is capable of reproducing with a drop in gain not to exceed 3 db may be found from Eqs. (9-33) and (9-34): $f_1 = 1/(2\pi 594{,}000 \times 0.01 \times 10^{-6}) = 26.8$ cycles/sec [where 594,000 is the value of R_e as found from Eqs. (9-25) and (9-11)], $f_2 = 1/(2\pi 79{,}000 \times 21 \times 10^{-12}) = 96{,}000$ cycles/sec.

The validity of neglecting the second term of Eq. (9-9) may now be checked. For resistance load $\cos \psi = 1$ so the equation becomes $C_g = 4 + 118.5 \times 0.005 = 4.6$ $\mu\mu f$. This indicates that C_T should have been 21.6 instead of 21, in the mid-frequency range, an inappreciable error.

If the frequency range is to be extended to 2 Mc, the proper value of R_1 may be found by substituting $C_T = 21 \times 10^{-12}$ and $f_2 = 2 \times 10^6$ into Eq. (9-34) and solving for R_0 to find $R_0 = 3800$ ohms. R_1 may now be found from Eq. (9-10) but it may be seen that for all practical purposes $R_1 = R_0$ when R_0 is so very small as compared to R_g and r_p. The mid-frequency gain of this wide-band amplifier may be found from Eq. (9-16) to be $A_m = 1500 \times 10^{-6} \times 3800 = 5.7$. This is a very low gain, and it is evident that the use of tubes with g_m of the order of 8000 to 10,000 μmhos is highly desirable. Incidentally, while decreasing R_1 improves the h-f gain, it causes a slight loss in the l-f response. Equation (9-33) now gives $f_1 = 1/(2\pi 503{,}000 \times 0.01 \times 10^{-6}) = 31.6$ cycles/sec.

✓**Practical Circuits.** Figure 9-8 shows circuits of practical resistance-coupled amplifiers with a single source for all direct potentials, usually a vacuum-tube rectifier such as described in Chap. 7. The resistance R_k provides grid bias, since the space current, in flowing through this resistance, produces a drop that is negative toward the ground end. The capacitance C_k is a by-pass condenser to prevent the alternating components of the plate current from setting up a voltage across the bias resistor R_k which would then be impressed on the grid along with the regular input signal.[1] The resistance R_d in Fig. 9-8b drops the plate supply voltage to the lower potential desired for the screen grid and also, together with the by-pass condenser C_d, prevents any alternating voltage from being impressed on the screen.

Effect of Screen-grid By-pass Condenser. Equation (9-30) was developed for a triode but may be used equally well for a pentode if the screen- and suppressor-grid potentials are main-

[1] This phenomenon, known as *feedback*, is sometimes desirable (see p. 361).

tained constant. Under these conditions the only alternating potentials are those applied to the control grid and the plate, exactly as for the triode.

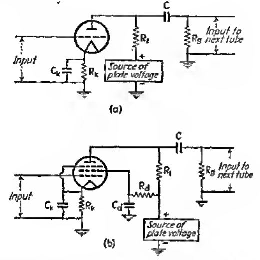


FIG. 9-8. Practical circuits of resistance-coupled amplifiers: (a) triode, (b) pentode.

In a practical circuit, such as that of Fig. 9-8b, some alternating potential may exist between the screen grid and cathode, owing to failure of the by-pass condenser C_d to present a sufficiently low impedance at low frequencies, and an additional l-f multiplying factor must be applied to Eq. (9-30). To determine this factor we may first consider the screen grid as the plate of a triode, with a voltage induced across its load circuit (consisting of C_d and R_d in parallel) due to the action of the control grid on the space current. Having determined this voltage, we may next consider the screen grid as a control grid, whence it will be found to act on the plate current in such a manner as to oppose the action of the regular control grid and so reduce the voltage that would otherwise be set up in the plate load circuit. Thus ineffective by-passing of the screen grid to cathode will cause a loss in gain at low frequencies which is analyzed in the following paragraphs.

To determine the alternating potential set up between screen grid and cathode, consider the equivalent circuit of Fig. 9-9 wherein the screen grid is treated as the plate of a triode. In this circuit μ_{sgs} is the mu factor of the screen grid relative to the control grid with the screen-grid current held constant (see definition of mu factor in Appendix A, page 604) and r_{g2} is the dynamic resistance of the screen grid, comparable to r_p for the plate, and equal to $\partial e_{c2}/\partial i_{c2}$ where e_{c2} and i_{c2} are the total instantaneous values of screen-grid voltage and current respectively (see Appendix A, page 609) The impedance Z_s is the impedance found between the screen grid and the cathode consisting essentially of the reactance of the condenser C_d in parallel with the resistance R_d. It is true that, in the circuit of Fig. 9-8b, the parallel combination of C_k and R_k is also in series, but the capacitance of C_k is normally so large compared to that of C_d that its effect on the screen-grid load circuit may be neglected without serious error. From the circuit of Fig. 9-9 we may evidently write

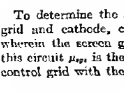


FIG. 9-9. Equivalent screen-grid circuit of a pentode tube.

$$\mathbf{E}_{g2} = -\mu_{sgs}\mathbf{E}_{g1}\frac{\mathbf{Z}_s}{r_{g2} + \mathbf{Z}_s} \tag{9-38}$$

which gives the alternating voltage between screen grid and cathode.

To determine the effect of this voltage on the plate circuit, we may draw the equivalent circuit of Fig. 9-10 showing the equivalent plate-circuit voltages of both the control and the screen grids. This may be compared with Fig. 3-24 for a triode. The symbol μ_{psp} is the mu factor of the plate with respect to the screen grid, with the plate current held constant.

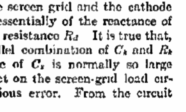


FIG. 9-10. Equivalent plate circuit of a pentode tube.

By inspection of Fig. 9-10 we may write

$$(-\mu\mathbf{E}_{g1}) + (-\mu_{psp}\mathbf{E}_{g2}) = \mathbf{I}_p(r_p + \mathbf{Z}_L) \tag{9-39}$$

Substituting Eq. (9-38) into (9-39) gives

$$(-\mu \mathbf{E}_{g1}) + \left(\mu_{psp}\mu_{sgs}\mathbf{E}_{g1}\frac{\mathbf{Z}_s}{r_{g2} + \mathbf{Z}_s}\right) = \mathbf{I}_p(r_p + Z_L) \qquad (9\text{-}40)$$

But, $\mu_{psp}\mu_{sgs} = \mu$ approximately[1] and, therefore, Eq. (9-40) reduces to

$$-\mu \mathbf{E}_{g1}\left(1 - \frac{\mathbf{Z}_s}{r_{g2} + \mathbf{Z}_s}\right) = \mathbf{I}_p(r_p + \mathbf{Z}_L) \qquad (9\text{-}41)$$

[1] This may be proved by writing equations for the screen-grid and plate currents similar to Eq. (3-9). Thus

$$i_{c2} = a_1 e_{c1} + b_1 e_{c2} + c_1 e_b + d_1 \qquad (1)$$

and differentiating partially with respect to e_{c1}, with i_{c2} and e_b constant, gives

$$0 = a_1 + b_1 \frac{\partial e_{c2}}{\partial e_{c1}} \qquad (2)$$

or

$$\frac{a_1}{b_1} = -\frac{\partial e_{c2}}{\partial e_{c1}} \qquad (3)$$

But, by definition (see Appendix A, page 604), $-\partial e_{c2}/\partial e_{c1}$ is the screen-grid–control-grid mu factor of the tube, μ_{sgs}. Therefore,

$$a_1/b_1 = \mu_{sgs} \qquad (4)$$

We may also write

$$i_b = a e_{c1} + b e_{c2} + c e_b + d \qquad (5)$$

where the factors a, b, c, d are, of course, different from a_1, b_1, c_1, d_1 used for Eq. (1).

Differentiating Eq. (5) partially with respect to e_{c1}, with i_b and e_b constant, gives

$$0 = a + b\frac{\partial e_{c2}}{\partial e_{c1}} \qquad (6)$$

or

$$\frac{a}{b} = -\frac{\partial e_{c2}}{\partial e_{c1}} = \mu_{sgp} \qquad (7)$$

A study of the characteristic curves of pentode tubes shows that the ratios a_1/b_1 and a/b are very nearly equal, so that we may write

$$\mu_{sgp} = \mu_{sgs} \qquad \text{(approx)} \qquad (8)$$

Differentiating Eq. (5) partially with respect to e_{c2}, with i_b and e_{c1} con-

or

$$-\mu \mathbf{E}_{g1} \frac{r_{g2}}{r_{g2} + \mathbf{Z}_s} = \mathbf{I}_p(r_p + \mathbf{Z}_L) \tag{9-42}$$

It is evident from Eq. (9-42) that the equivalent circuit of Fig. 9-10 may now be replaced by the circuit of Fig. 9-11 using only

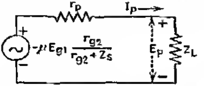


Fig. 9-11. Simplified equivalent circuit of a pentode tube.

one generator. It is further evident that the circuit of Fig. 9-5 may now be modified to take into account the effect of the screen-grid by-pass condenser by merely replacing the voltage $\mathbf{E}_s$ with a

---

stant, gives

$$0 = b + c \frac{\partial e_b}{\partial e_{c2}} \tag{9}$$

or

$$\frac{b}{c} = -\frac{\partial e_b}{\partial e_{c2}} = \mu_{psp} \tag{10}$$

and differentiating Eq. (5) partially with respect to e_{c1}, with i_b and e_{c2} constant, gives

$$0 = a + c \frac{\partial e_b}{\partial e_{c1}} \tag{11}$$

or

$$\frac{a}{c} = -\frac{\partial e_b}{\partial e_{c1}} = \mu \tag{12}$$

Multiplying Eqs. (7) and (10) and making the substitution of Eq. (8) gives

$$\frac{a}{b}\frac{b}{c} = \mu_{gsp}\mu_{psp} = \mu_{sg}\mu_{psp} \qquad \text{(approx)} \tag{13}$$

But the left-hand side of this equation reduces to a/c which, from Eq. (12), is μ. Therefore, Eq. (13) becomes

$$\mu = \mu_{sg}\mu_{psp} \qquad \text{(approx)} \tag{14}$$

which was to be proved.

new voltage $\mathbf{E}_g r_{g2}/(r_{g2} + \mathbf{Z}_s)$. Actually this change need be made only in the l-f circuit of Fig. 9-5c since the shunting effect of the condenser C_d is so effective in the mid-frequency band and at higher frequencies as to make $\mathbf{Z}_s$ equal to zero for all practical purposes. With this change a new equation may be developed for the l-f response as follows:

Replace $\mathbf{E}_g$ in Eq. (9-27) with the new value from Fig. 9-11.

$$\mathbf{E}_o' = -g_m \mathbf{E}_g \frac{r_{g2}}{r_{g2} + \mathbf{Z}_s} R_0 \frac{R_c}{R_c - jX_c} \tag{9-43}$$

Let

$$\mathbf{Z}_s = R_s - jX_s \tag{9-44}$$

Then Eq. (9-43) may be written

$$\mathbf{E}_o' = -g_m \mathbf{E}_g R_0 \frac{r_{g2}}{r_{g2} + R_s - jX_s} \frac{R_c}{R_c - jX_c} \tag{9-45}$$

and the l-f gain becomes, in polar form,

$$\mathbf{A}_l = \frac{\mathbf{E}_o'}{\mathbf{E}_g} = \underbrace{g_m R_0}_{\substack{\text{Mid-}\\ \text{frequency}\\ \text{gain}}} \underbrace{\frac{1}{\sqrt{\dfrac{(r_{g2} + R_s)^2}{r_{g2}^2} + \dfrac{X_s^2}{r_{g2}^2}}}}_{\substack{\text{Low-frequency}\\ \text{multiplying}\\ \text{factor for } C_d}} \underbrace{\frac{1}{\sqrt{1 + \dfrac{X_c^2}{R_c^2}}}}_{\substack{\text{Low-frequency}\\ \text{multiplying}\\ \text{factor for } C}}$$

$$\Big/ \underbrace{180°}_{\substack{\text{Tube phase}\\ \text{reversal}}} + \underbrace{\tan^{-1}\frac{X_s}{r_{g2} + R_s}}_{\substack{\text{Phase shift}\\ \text{due to } C_d}} + \underbrace{\tan^{-1}\frac{X_c}{R_c}}_{\substack{\text{Phase shift}\\ \text{due to } C}} \tag{9-46}^*$$

For most purposes Eq. (9-46) may be further simplified by assuming that R_d is much larger than the reactance of C_d. This will be true in a normal amplifier for all frequencies at which the gain is within a few decibels of the mid-frequency gain. Under this assumption we may write

$$\mathbf{Z}_s = -jX_d \tag{9-47}$$

where $X_d = 1/\omega C_d$. Substituting this value of $\mathbf{Z}_s$ into Eq. (9-43) we may develop an equation similar to (9-46) in which the l-f multiplying factor for C_d is

$$\frac{1}{\sqrt{1 + (X_d^2/r_{g2}^2)}}$$

A new equation may now be written giving the gain of the amplifier at any frequency by modifying Eq. (9-30) as follows:

$$\mathrm{A} = \underbrace{g_m R_0}_{\substack{\text{Mid-}\\\text{frequency}\\\text{gain}}} \underbrace{\frac{1}{\sqrt{1 + (R_0^2/X_T^2)}}}_{\substack{\text{High-frequency}\\\text{multiplying}\\\text{factor}}} \underbrace{\frac{1}{\sqrt{1 + (X_c^2/R_e^2)}}}_{\substack{\text{Low frequency}\\\text{multiplying}\\\text{factor for } C}} \underbrace{\frac{1}{\sqrt{1 + (X_d^2/r_{g2}^2)}}}_{\substack{\text{Low frequency}\\\text{multiplying}\\\text{factor for } C_d}}$$

$$\Big/ \underbrace{180°}_{\substack{\text{Tube phase}\\\text{reversal}}} - \underbrace{\tan^{-1}\frac{R_0}{X_T}}_{\substack{\text{High frequency}\\\text{phase shift}}} + \underbrace{\tan^{-1}\frac{X_c}{R_e}}_{\substack{\text{Low frequency}\\\text{phase shift}\\\text{due to } C}} + \underbrace{\tan^{-1}\frac{X_d}{r_{g2}}}_{\substack{\text{Low frequency}\\\text{phase shift}\\\text{due to } C_d}} \quad (9\text{-}48)^*$$

Equation (9-48) must not, of course, be used for such low frequencies that the approximation of Eq. (9-47) is invalid. For lower frequencies Eq. (9-46) must be used. Also this equation does not include the effect of C_k.

The curves of Fig. 9-7 may be used to evaluate the l-f multiplying factor for C_d in Eq. (9-48) by considering the abscissa to be a plot of f_3/f where f_3 is the frequency at which the multiplying factor has a value of 0.707, *i.e.*, the frequency at which $X_d = r_{g2}$.

Effect of Cathode By-pass Condenser. The cathode by-pass condenser C_k in both circuits of Fig. 9-8 will affect the l-f response in much the same manner as does C_d, except that the voltage across this condenser is a function of the plate (or output) current. Because the plate current is affected by C_T, C, and C_k, any expression correcting for the effect of C_k must be a function of all three capacitances. This complicates the solution, which is omitted here but may be worked out as a problem in feedback (see page 301 for analysis of feedback in amplifiers).

✓**Minimum Frequency Response of Pentode Amplifiers.** Equations (9-31) and (9-33) for the minimum frequency response of an amplifier are not valid for the amplifier circuits of Fig. 9-8 since they do not include the effects of condensers C_d and C_k. It would be possible to determine a new equation for the minimum frequency which included the effect of C_d by solving for f from the last *two* factors of Eq. (9-48), but the resulting equation would not be simple in form. Probably a better procedure is to solve for f_1 from Eq. (9-33) and then, using this frequency, compute the multiplying factor for C_d from Eq. (9-48). If this factor is appreciably less than unity, upward revision of the estimated value of

f_1 can be made until Eq. (9-48) gives a value for A which is 70.7 per cent of the mid-frequency gain.

Example. Consider the amplifier of the example on page 277 with $R_1 =$ 100,000 ohms. Let $r_{g2} = 40{,}000$ ohms, $R_d = 1{,}000{,}000$ ohms, $C_d = 0.1\ \mu f$. Putting these values into the last term of Eq. (9-48) gives 0.55 for the multiplying factor due to C_d at the frequency $f_1 = 26.8$ cycles/sec computed from Eq. (9-33) in the original example. Since this value of f_1 was predicated on a drop in gain to 70.7 per cent due to the effect of condenser C, the actual gain, considering both C and C_d, is $0.55 \times 0.707 = 0.39$. We must now find the frequency at which the product of both the l-f multiplying factors of Eq. (9-48) equals 0.707. For small changes in frequency, the gain of the amplifier increases almost in proportion to the frequency. Let us therefore guess that the correct minimum frequency is approximately $26.8 \times (0.707/0.39) = 48.5$, and then round this figure out to 50 since the gain does not increase quite in proportion to the frequency. Substituting numerical values into the two l-f multiplying factors of Eq. (9-48) for $f = 50$ gives 0.88 for the factor for C and 0.78 for the factor for C_d. The product of these two is approximately 0.7. Therefore, the minimum frequency of the amplifier is approximately 50 cycles for a drop in gain of 3 db from the mid-frequency gain, assuming that C_k is sufficiently large to make its impedance negligible.

Multistage Amplifiers. As pointed out in a preceding section, video amplifiers must be built with a relatively low gain per stage to achieve the desired h-f response. The required over-all gain is then obtained by adding a sufficient number of additional stages of amplification. This indicates that the band width of a multistage amplifier may be increased indefinitely without affecting the over-all gain by reducing the gain per stage and increasing the number of stages. However if this procedure is carried too far, the gain per stage will eventually become less than unity and no amplification will be possible. Thus there is a limit to the band width obtainable in this manner.[1]

Another problem in designing multistage, wide-band amplifiers is that, for a fixed band width, the flatness of each stage must be increased as the number of stages is increased. If the gain of the complete amplifier must be kept within 3 db of the mid-frequency gain and there are n stages, it is obvious that the average drop in gain of each stage must not exceed $3/n$ db at the minimum and maximum frequencies to be amplified. A similar problem is

[1] For an analysis of this problem, see W. R. McLean, Ultimate Bandwidths in High-gain Multistage Video Amplifiers, *Proc. IRE*, 32, p. 12, January, 1944.

encountered in keeping the total phase shift of the complete amplifier within the required limits. This imposes very severe requirements on the design of amplifiers that employ a large number of stages.

To meet the foregoing problems video amplifiers are often designed with special l-f and h-f compensating circuits, some of which are described in succeeding sections of this chapter.

Low-frequency Compensation. *Figure* 9-12 shows a simple and effective circuit for improving the l-f response of a video amplifier. R_1 is the conventional coupling resistance of the resistance-coupled amplifier of Fig. 9-8*b*, but a condenser C_2 is connected in series with R_1, with a resistance R_2 providing a d-c path for the plate current. The general effect of this combination is that the condenser C_2 presents negligible impedance in the mid-

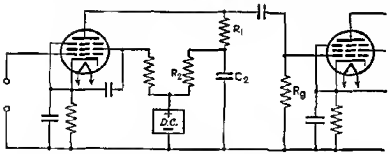


FIG. 9-12. Low-frequency compensating circuit.

frequency region, and the amplifier performance is exactly as described in preceding sections. As the frequency becomes lower and the coupling condenser causes a falling off in gain, the reactance of C_2 increases and so provides an increase in the total coupling impedance. With proper design this increase is sufficient to offset the effect of C down to a rather low frequency. At the same time the phase shift produced by C_2 is opposite in direction to that produced by C, and proper design will permit operation at very much lower frequencies with only a small phase shift.

The equivalent circuit for Fig. 9-12 is shown in Fig. 9-13*a* for low frequencies, where Z_2 represents the impedance of R_2 and C_2 in parallel. By applying Norton's theorem at points AB and by noting that $(R_1 + Z_2) \ll r_p$, it may be simplified to that of Fig. 9-13*b*.

Solution of the circuit of Fig. 9-13*b* requires evaluation of the shunting impedance, $R_1 + Z_2$. From Fig. 9-12 we may write

$$R_1 + Z_2 = R_1 + \frac{1}{(1/R_2) - (1/jX_2)} \qquad (9\text{-}49)$$

assuming negligible impedance in the d-c source and letting $X_2 = 1/\omega C_2$. If the denominator of Eq. (9-49) is rationalized, we get

$$R_1 + Z_2 = \frac{R_1(R_2^2 + X_2^2) + R_2X_2^2 - jR_2^2X_2}{R_2^2 + X_2^2} \qquad (9\text{-}50)$$

At all frequencies for which the circuit of Fig. 9-12 provides adequate compensation, R_2 is appreciably larger than X_2, so that we may write $X_2^2 \ll R_2^2$. We may, therefore, simplify Eq.

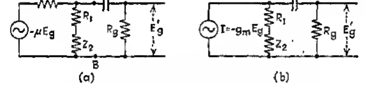


FIG. 9-13. Equivalent circuits for Fig. 9-12.

(9-50) by neglecting X_2^2 in the denominator and in the parentheses in the numerator, giving

$$R_1 + Z_2 = \frac{R_1R_2^2 + R_2X_2^2 - jR_2^2X_2}{R_2^2} = R_1 + \frac{X_2^2}{R_2} - jX_2 \qquad (9\text{-}51)$$

We may now solve for the gain by following the same procedure as in Eqs. (9-21) to (9-24). Comparison of Fig. 9-13*b* with Fig. 9-5*e* shows that R_a in Eqs. (9-21) to (9-24) now becomes $R_1 + Z_2$ and all other terms remain the same. Equation (9-24) may, therefore, be written, for the compensated amplifier,

$$E_g' = -g_m E_g \frac{\left(R_1 + \dfrac{X_2^2}{R_2} - jX_2\right)R_g}{R_1 + \dfrac{X_2^2}{R_2} - jX_2 + R_g - jX_c} \qquad (9\text{-}52)$$

In a video amplifier $R_1 \ll R_g$; therefore, we may neglect R_1 in the denominator. The gain may, therefore, be written, in polar form, as

$$\mathrm{A}_l = \frac{\mathrm{E}_o'}{\mathrm{E}_g} = g_m R_l \frac{\sqrt{\left(R_1 + \frac{X_2^2}{R_2}\right)^2 + X_2^2} \;\Big/ 180^\circ - \tan^{-1}\frac{X_2 R_2}{R_1 R_2 + X_2^2}}{\sqrt{\left(R_g + \frac{X_2^2}{R_2}\right)^2 + (X_2 + X_c)^2} \;\Big/ -\tan^{-1}\frac{(X_2 + X_c)R_2}{R_2 R_g + X_2^2}} \tag{9-53}$$

In the denominator X_2^2/R_2 is negligible compared to R_g.[1] Equation (9-53) may therefore be simplified and rearranged to give

$$\mathrm{A}_l = \underbrace{g_m R_l}_{\text{Mid frequency gain}} \underbrace{\sqrt{\left(1 + \frac{X_2^2}{R_1 R_2}\right)^2 + \frac{X_2^2}{R_1^2}}}_{\text{Low frequency compensating factor}} \underbrace{\frac{1}{\sqrt{1 + \frac{(X_2 + X_c)^2}{R_g^2}}}}_{\text{Low frequency multiplying factor}}$$

$$\Big/ \underbrace{180}_{\text{Tube phase reversal}} - \underbrace{\tan^{-1}\frac{X_2 R_2}{R_1 R_2 + X_2^2}}_{\text{Compensating phase shift}} + \underbrace{\tan^{-1}\frac{X_2 + X_c}{R_g}}_{\text{Low frequency phase shift}} \tag{9-54)*$$

Equation (9-54) does not include the effect of condensers C_d and C_k of Fig. 9-8 although the effect of C_d may be included by inserting the multiplying factor and phase shift from Eq. (9-40).

As previously stated, perfect phase compensation will be obtained if the angle of A varies in direct proportion to the frequency or if it is zero at all frequencies. Thus Eq. (9-54) shows that perfect phase compensation will be obtained with the circuit of Fig. 9-12 if $X_2R_2/(R_1R_2 + X_2^2) = (X_2 + X_c)/R_g$. This is an impossible condition to realize at all frequencies since the right-hand side of this equation varies inversely with frequency while the left-hand side varies in a complex manner, but R_2 is usually of

[1] The purpose of the condenser C_2, Fig. 9-12, is to increase the impedance of the load circuit slightly at low frequencies. To do so, its reactance must be of the same order of magnitude as the resistance of R_1, since an appreciably lower reactance would have negligible effect whereas a higher reactance would make the load impedance vary almost inversely as the frequency and so would overcompensate. The resistance R_2 is usually large enough so that it has little effect on the frequency characteristic, its primary function being to provide a d-c path for the plate current, and is therefore larger than R_1. Since R_1 is much smaller than R_g, it follows that $X_2^2/R_2 \ll R_g$.

such size that $X_2^2 \ll R_1R_2$ at all except the very lowest frequencies so that $X_2R_2/(R_1R_2 + X_2^2) \cong X_2/R_1$. Under these circumstances nearly perfect phase compensation will be obtained if $X_2/R_1 = (X_c + X_2)/R_g$, a relationship that is physically realizable. If we rewrite Eq. (9-54) by letting $X_2^2 \ll R_1R_2$ and if the multiplying factor and phase shift for C_d are also included, the result is

$$\mathrm{A}_l = \underbrace{g_m R_1}_{\substack{\text{Mid-}\\\text{frequency}\\\text{gain}}} \underbrace{\sqrt{1 + \frac{X_2^2}{R_1^2}}}_{\substack{\text{Low-frequency}\\\text{compensating}\\\text{factor}}} \underbrace{\frac{1}{\sqrt{1 + \dfrac{(X_2 + X_c)^2}{R_g^2}}}}_{\substack{\text{Low-frequency}\\\text{multiplying}\\\text{factor due}\\\text{to } C}} \underbrace{\frac{1}{\sqrt{1 + \dfrac{X_d^2}{r_{g2}^2}}}}_{\substack{\text{Low-frequency}\\\text{multiplying}\\\text{factor due}\\\text{to } C_d}}$$

$$\Big/ \underbrace{180°}_{\substack{\text{Tube phase}\\\text{reversal}}} \underbrace{- \tan^{-1} \frac{X_2}{R_1}}_{\substack{\text{Compensating}\\\text{phase shift}}} \underbrace{+ \tan^{-1} \frac{X_2 + X_c}{R_g}}_{\substack{\text{Phase shift}\\\text{due to } C}} \underbrace{+ \tan^{-1} \frac{X_d}{r_{g2}}}_{\substack{\text{Phase shift}\\\text{due to } C_d}} \qquad (9\text{-}55)^*$$

Equation (9-55) gives the gain of a resistance-coupled, low-frequency-compensated amplifier at low frequencies assuming that $X_2^2 \ll R_2^2$, $X_2^2 \ll R_1R_2$, $R_1 \ll R_g$, and that the impedance of C_k is negligible. With the possible exception of the next to the last, these are all conditions that are normally realized in a video amplifier within the desired frequency range.

From Eq. (9-55) it is evident that perfect phase compensation, including the effect of C_d, will be obtained if

$$\tan^{-1} \frac{X_2 + X_c}{R_g} + \tan^{-1} \frac{X_d}{r_{g2}} - \tan^{-1} \frac{X_2}{R_1} = 0 \qquad (9\text{-}56)$$

This will be approximately true for total phase shifts up to about 30 deg in the uncompensated amplifier if[1]

$$\frac{X_2 + X_c}{R_g} + \frac{X_d}{r_{g2}} = \frac{X_2}{R_1} \qquad (9\text{-}57)$$

If we rewrite the reactances of Eq. (9-57) in terms of ω, (*i.e.*, $X_c = 1/\omega C$, $X_2 = 1/\omega C_2$, $X_d = 1/\omega C_d$), we may solve Eq. (9-57) for C_2 to obtain

[1] The tangent and its angle are approximately equal for angles of 30 deg or less.

$$C_2 = \frac{CC_d}{CR_g + C_d r_{g2}} \frac{r_{g2}(R_g - R_1)}{R_1} \tag{9-58}$$

Since $R_1 \ll R_g$, this may be simplified to

$$C_2 = \frac{CC_d}{CR_g + C_d r_{g2}} \cdot \frac{r_{g2} R_g}{R_1} \tag{9-59}$$

Further inspection of Eq. (9-55) will disclose that the relationship of Eq. (9-57) makes the compensation multiplying factor *nearly equal to the reciprocal of the product of the two l-f multiplying factors*. Thus excellent compensation will be obtained down to the lowest frequency for which the approximations used in deriving Eq. (9-55) are valid. In a normal amplifier, satisfactory amplitude compensation can be obtained to a lower fre-

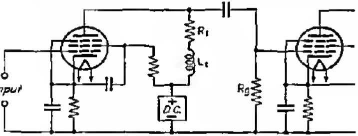


FIG. 9-14. Shunt-peaking, h-f compensating circuit.

quency than can satisfactory phase compensation. Thus the lowest frequency at which satisfactory compensation can be obtained is normally that which makes X_2 approximately one-half R_1, giving a phase-shift compensation of about 30 deg. At still lower frequencies compensation becomes increasingly less perfect, and phase shift and, to a lesser extent, gain will deviate considerably from mid-frequency values.

High-frequency Compensation. Figure 9-14 shows a simple but effective circuit for compensating the h-f response of a video amplifier. This method of compensation is commonly known as *shunt peaking*. The inductance L_1 is of such size that it has negligible effect in the mid-frequency range but causes the impedance of the L_1, R_1 circuit to increase at frequencies where the reactance of C_T becomes low, thus tending to hold the gain constant at higher frequencies. The phase shift due to L_1 is opposite to that

produced by C_T so that compensation for phase as well as magnitude is provided. (C_T was defined in Fig. 9-5, page 266.)

The equivalent circuit at high frequencies is shown in Fig. 9-15*a* which may be reduced to that of Fig. 9-15*b* by the use of Norton's theorem applied at AB and by noting that $(R_1 + j\omega L_1) \ll R_g < r_p$ in amplifiers which require h-f compensation (such as video amplifiers). Evidently,

$$\mathbf{E}_g' = \mathbf{I}\,\frac{1}{\dfrac{1}{R_1 + jX_1} - \dfrac{1}{jX_T}} = -g_m\mathbf{E}_g\,\frac{jX_T(R_1 + jX_1)}{jX_T - R_1 - jX_1} \tag{9-60}$$

where $X_1 = \omega L_1$. Dividing numerator and denominator by j and dividing both sides of the equation by $\mathbf{E}_g$ gives

$$\mathbf{A}_h = \frac{\mathbf{E}_g'}{\mathbf{E}_g} = -g_m X_T\,\frac{R_1 + jX_1}{(X_T - X_1) + jR_1} \tag{9-61}$$

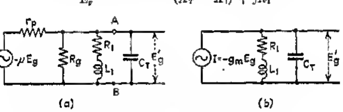


FIG. 9-15. Equivalent circuits for Fig. 9-14.

In polar form

$$\mathbf{A}_h = g_m X_T\,\frac{\sqrt{R_1^2 + X_1^2}\;\angle 180° + \tan^{-1}(X_1/R_1)}{\sqrt{(X_T - X_1)^2 + R_1^2}\;\angle \tan^{-1}\dfrac{R_1}{X_T - X_1}} \tag{9-62}$$

or, dividing numerator and denominator by X_T and taking R_1 out of the radical

$$\mathbf{A}_h = \underbrace{g_m R_1}_{\substack{\text{Mid-}\\ \text{frequency}\\ \text{gain}}}\;\underbrace{\sqrt{1 + \frac{X_1^2}{R_1^2}}}_{\substack{\text{Compensating}\\ \text{multiplying}\\ \text{factor}}}\;\underbrace{\frac{1}{\sqrt{\dfrac{(X_T - X_1)^2}{X_T^2} + \dfrac{R_1^2}{X_T^2}}}}_{\substack{\text{High-frequency}\\ \text{multiplying}\\ \text{factor}}}$$

$$\angle\,\underbrace{180°}_{\substack{\text{Tube phase}\\ \text{reversal}}} + \underbrace{\tan^{-1}\frac{X_1}{R_1}}_{\substack{\text{Compensating}\\ \text{phase shift}}} - \underbrace{\tan^{-1}\frac{R_1}{X_T - X_1}}_{\substack{\text{High-frequency}\\ \text{phase shift}}} \tag{9-63}*$$

Equation (9-63) gives the gain of the compensated amplifier of Fig. 9-14 at all frequencies in the mid-frequency band and above, under the conditions that $(R_1 + jX_1)$ is very much smaller than either R_g or r_p. A study of this equation will show that the first radical following $g_m R_1$ is a term which increases above unity in magnitude as the frequency increases and thus provides h-f compensation, while the second radical is merely a modification of the h-f multiplying factor of Eq. (9-20). That this is so may be seen by letting $X_1 = 0$, as in the uncompensated amplifier, whence this radical will reduce to the h-f multiplying factor of

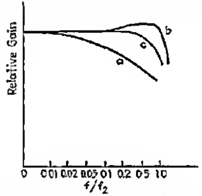


FIG. 9-16. Curves of h-f response. (*a*) no compensation, (*b*) compensation with $2\pi f_2 L/R_1 = 0.5$, (*c*) compensation with $2\pi f_2 L/R_1 = 0.414$.

Eq. (9-20), since we assumed that $R_1 = R_0$ in the compensated amplifier.

Equation (9-63) does not lend itself to analysis so readily as did the equations for l-f compensation. However if we let $Q = 2\pi f_2 L_1/R_1$, where f_2 is the frequency at which the reactance of C_T is equal to the resistance of R_1 as given by Eq. (9-34), we may use Eq. (9-63) to compute data with which to plot a series of curves showing gain as a function of frequency for various values of Q. Three such curves are shown in Fig. 9-16. Curve *b* for $Q = 0.5$ produces perfect compensation at $f = f_2$ but causes a slight increase in gain at lower frequencies. The maximum gain occurs at approximately $f = 0.7f_2$ and is equal to 1.03 times the mid-frequency gain. In a single stage of amplification this increase is unimportant, but it may be excessive in a 10- or 20-stage amplifier.

Curve c for $Q = 0.414$ is the flattest curve obtainable. Both curves should be compared with the uncompensated curve a ($Q = 0$).

Zero phase shift cannot be obtained with this circuit as may be seen if an attempt is made to solve for a value of X_1 which will make the angle of Eq. (9-63) equal to 180 deg, but there is appreciable improvement in the phase characteristic with the best results at about $Q = 0.35$. The optimum value of Q for phase compensation, therefore, does not correspond to the optimum value for amplitude compensation, and the usual procedure is to make Q somewhere between 0.4 and 0.5.

Both h-f and l-f compensation may be combined in a single circuit by adding L_1 in series with R_1 in the circuit of Fig. 9-12. Equations (9-55) and (9-63) may be readily combined to give a single equation for the fully compensated amplifier which is applicable at all frequencies.

Other Circuits for High-frequency Compensation. There are a number of other h-f compensating circuits most of which include an inductance connected in series with the coupling condenser C and are, therefore, known as *series-peaking* circuits. Such an arrangement separates the equivalent capacitance C_T into two parts, one being the output capacitance of the driving tube together with the associated wiring capacitance and the other being the input capacitance of the driven tube together with wiring capacitance. This arrangement decreases the effectiveness of these capacitances in reducing the h-f gain of the amplifier and thus produces improved compensation with respect to maintaining a flat frequency response characteristic. Unfortunately the transient response of such circuits is not too satisfactory, and they are more difficult to adjust than the shunt-peaking circuit of Fig. 9-14.

Circuits of the series type may be analyzed by treating them as low-pass filters and then applying the equations of such filters to the solution of the amplifier circuit. By extending this approach it is possible to improve further the flatness of the gain characteristic by incorporating m-derived filter sections into the tube coupling network.[1]

[1] For details of this method of approach, see Austin V. Eastman, The Application of Filter Theory to the Design of Reactance Networks, *Proc. IRE*, **32**, p. 538, September, 1944.

The performance of the shunt-peaking circuit of Fig. 9-14 may be improved by inserting a suitable condenser in parallel with L_1. This permits the inductance to be made sufficiently small to avoid the rise in gain of curve *b*, Fig. 9-16, while holding the gain constant to a somewhat higher frequency than in curve *c*. Thus the wider band width of curve *b* is obtained without a rise in gain below f_1. With proper design the condenser to be used across the coil may be the distributed capacitance of the coil winding.[1]

Transformer-coupled Voltage Amplifiers. Transformer coupling is sometimes used in a circuit such as that of Fig. 9-17, the equivalent circuit of which is shown in Fig. 9-18. R_1 and R_2

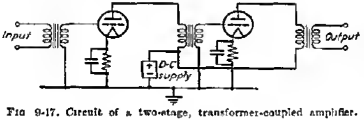


FIG. 9-17. Circuit of a two-stage, transformer-coupled amplifier.

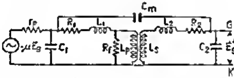


FIG. 9-18. Complete equivalent circuit of a transformer-coupled amplifier

are the resistances of the primary and secondary windings respectively, and L_1 and L_2 represent the leakage reactances. C_1 represents the distributed capacitance of the primary winding together with the plate-to-cathode capacitance of the driving tube and any capacitances in the wiring between the tube and the transformer. Similarly C_2 represents the distributed capacitance of the secondary winding, the grid-to-cathode capacitance of the driven tube, and any wiring capacitance on the secondary side of the transformer. C_m represents capacitance between the two transformer

[1] Alexander B. Bereskin, Improved High-frequency Compensation for Wide-band Amplifiers, *Proc IRE*, **32**, p. 608, October, 1944.

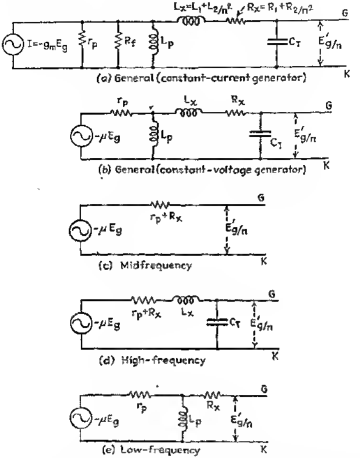


FIG. 9-19. Simplified equivalent circuits of a transformer-coupled amplifier.

windings. R_f represents iron losses and will actually vary somewhat with frequency, although the performance of the transformer may be predicted with sufficient accuracy for most purposes by neglecting its effect entirely.

Figure 9-19a shows the equivalent circuit of the transformer reduced to unity turns ratio with some simplification in the constants and with the constant-voltage generator $-\mu E_g$ replaced by

a constant-current generator $-g_m\mathbf{E}_g$ through application of Norton's theorem to the circuit lying to the left of C_1 in Fig. 9-18. (See discussion of Norton's theorem on page 267.) The primary resistance and leakage reactance are combined with the secondary resistance and reactance on the output side of the no-load inductance L_p, since the primary resistance and leakage reactance are normally too small compared with R_f and L_p for this arrangement to introduce appreciable error. The secondary resistance and reactance are reduced by the square of the transformer turn ratio. The capacitance C_T is made up of the secondary capacitance C_2, increased by the square of the turns ratio, and the mutual capacitance C_m, multiplied by a factor which is either a little larger or a little smaller than the square of the turns ratio, depending on whether the relative polarities of the transformer windings produce a voltage across C_m equal to the sum or to the difference of the secondary and primary voltages. Capacitance C is omitted entirely since its effect is normally too small to be of importance. The resistance R_f is normally so much larger than r_p in a triode that it may be neglected in triode amplifiers. This is not true for pentodes but transformer coupling is rarely used with pentode tubes except in the output circuit of a power amplifier where the problems are quite different (see material beginning on page 357). Therefore R_f will be omitted from the remainder of this discussion on transformer-coupled voltage amplifiers. With this simplification, Thévenin's theorem may be used to change the constant-current circuit of Fig. 9-19a back to the constant-voltage circuit of Fig. 9-19b for analysis of the amplifier performance in each of three different frequency ranges.

In the middle range of frequencies the capacitances and inductances have negligible effect, and the circuit reduces to that of Fig. 9-19c. Since no current is flowing in this circuit, the gain will be constant throughout the range of frequencies to which the circuit applies.

At higher frequencies the leakage reactance and shunting capacitance must be considered and the circuit of Fig. 9-19d must be used. It should be apparent from this circuit that there is a frequency at which L_x and C_T are in series resonance, tending to raise the output voltage above its mid-frequency value. The magnitude of this effect is dependent largely upon the series resistance $r_p + R_x$, since a low series resistance will permit a comparatively large

current to flow through the series-resonant circuit, producing a very high output voltage across the condenser, whereas a sufficiently high resistance will entirely eliminate the resonance effect. Evidently if the frequency is increased above the resonance point, the decreasing reactance of C_T will cause the output voltage to approach zero.

At frequencies below the middle range, the no-load reactance of the primary (ωL_p) becomes sufficiently low to affect the gain; *i.e.*, the transformer charging current becomes the major portion of the plate current. The circuit of the amplifier then reduces to that of Fig. 9-19c. Here it is evident that the gain begins to fall off at a frequency that makes the reactance of the coil L_p of the order of magnitude of r_p. As the frequency is further decreased, the gain continues to fall, approaching zero

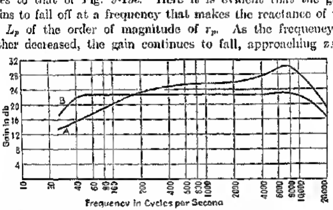


FIG. 9-20. Gain vs. frequency characteristic of two transformer-coupled amplifiers; the amplifier of curve A using an older type of transformer and that of curve B a more modern transformer.

The foregoing effects are fully illustrated in Fig. 9-20, curve A, where both the falling off in gain at low frequencies due to the effect of L_p and the resonant rise in voltage at the higher frequencies are plainly evident. Curve B represents a transformer of improved design, and the effects, although present, are much less in magnitude.

Gain of a Transformer-coupled Amplifier. The gain of a transformer-coupled amplifier in the middle range of frequencies may be determined directly from the circuit of Fig. 9-19c by noting that no current is flowing through the circuit and, therefore,

$$\frac{\mathbf{E}_g'}{n} = -\mu \mathbf{E}_g \tag{9-64}$$

The mid-frequency gain is then

$$\mathbf{A}_m = \frac{\mathbf{E}_g'}{\mathbf{E}_g} = -\mu n = \mu n \underline{/180^\circ} \qquad (9\text{-}65)^*$$

At the higher frequencies the gain must be found from Fig. 9-19*d*. Since this is a simple series circuit, the voltages E_g'/n and $-\mu \mathbf{E}_g$ are proportional to the respective impedances across which each appears or

$$\frac{\mathbf{E}_g'/n}{-\mu \mathbf{E}_g} = \frac{-jX_T}{(r_p + R_s) - j(X_T - X_s)} \qquad (9\text{-}66)$$

where $X_T = 1/\omega C_T$ and $X_s = \omega L_s$. Dividing numerator and denominator by $-j$, the gain may be written

$$\mathbf{A}_h = \frac{\mathbf{E}_g'}{\mathbf{E}_g} = \frac{-\mu n X_T}{(X_T - X_s) + j(r_p + R_s)} \qquad (9\text{-}67)$$

Dividing numerator and denominator by X_T and changing to polar form gives

$$\mathbf{A}_h = \mu n \frac{1}{\sqrt{\dfrac{(X_T - X_s)^2}{X_T^2} + \dfrac{(r_p + R_s)^2}{X_T^2}}} \underline{\Big/ 180^\circ - \tan^{-1} \frac{r_p + R_s}{X_T - X_s}} \qquad (9\text{-}68)^*$$

At the lower frequencies the circuit of Fig. 9-19*c* must be used. Here again $\mathbf{E}_g'/n$ and $-\mu \mathbf{E}_g$ are proportional to the impedances or

$$\frac{\mathbf{E}_g'/n}{-\mu \mathbf{E}_g} = \frac{jX_p}{r_p + jX_p} \qquad (9\text{-}69)$$

where $X_p = \omega L_p$. The gain may, therefore, be written

$$\mathbf{A}_l = \frac{\mathbf{E}_g'}{\mathbf{E}_g} = \frac{-\mu n X_p}{X_p - jr_p} \qquad (9\text{-}70)$$

Dividing numerator and denominator by X_p and converting to polar form gives

$$\mathbf{A}_l = \mu n \frac{1}{\sqrt{1 + \dfrac{r_p^2}{X_p^2}}} \underline{\Big/ 180^\circ + \tan^{-1} \frac{r_p}{X_p}} \qquad (9\text{-}71)^*$$

Equations (9-65), (9-68), and (9-71) may now be combined to give a single equation which expresses the performance of the normal, transformer-coupled, voltage amplifier using triode tubes.

$$A = \underbrace{\mu n}_{\substack{\text{Mid-}\\\text{frequency}\\\text{gain}}} \underbrace{\frac{1}{\sqrt{\dfrac{(X_T - X_x)^2}{X_T^2} + \dfrac{(r_p + R_x)^2}{X_T^2}}}}_{\substack{\text{High-frequency}\\\text{multiplying}\\\text{factor}}} \underbrace{\frac{1}{\sqrt{1 + \dfrac{r_p^2}{X_p^2}}}}_{\substack{\text{Low-frequency}\\\text{multiplying}\\\text{factor}}}$$

$$\Big/ \underbrace{180°}_{\substack{\text{Tube phase}\\\text{reversal}}} - \underbrace{\tan^{-1}\frac{r_p + R_x}{X_T - X_x}}_{\substack{\text{High-frequency}\\\text{phase shift}}} + \underbrace{\tan^{-1}\frac{r_p}{X_p}}_{\substack{\text{Low-frequency}\\\text{phase shift}}} \qquad (9\text{-}72)^*$$

While Eq. (9-72) was developed for triodes, it may be adapted for use with pentodes by replacing r_p with R_0, and μ with $g_m R_0$ (since μ was originally introduced from the product $g_m r_p$ in changing from a to b in Fig. 9-19), where R_0 is the parallel resistance of r_p and R_f and is nearly equal to R_f. These statements may be readily verified by repeating the preceding demonstration without neglecting R_f in the circuits of Fig. 9-19.

Actually pentodes are rarely used with transformer-coupled amplifiers for reasons that will become evident from the discussion in the next section.

Frequency Response of a Transformer-coupled Amplifier. Triode tubes used with transformer-coupled, voltage amplifiers are usually of the general-purpose variety having a μ that is not so high as the high-μ type used in resistance-coupled amplifiers or so low as the low-μ tubes used in power amplifiers. A μ of 8 to 20 is typical, with a plate resistance of perhaps 7500 to 15,000 ohms.

The desirability of avoiding high-μ tubes (and, therefore, tubes with high r_p) may be seen by considering the l-f multiplying factor of Eq. (9-72). To secure good l-f response this fraction must approach unity; *i.e.*, X_p must be large or r_p must be small. An increase in X_p may be obtained by (1) increasing the number of turns on the transformer primary, (2) increasing the cross section of the core, and (3) using a core material of higher permeability, where the incremental permeability must be considered because of the direct component of the plate current flowing through the primary winding (see page 192). Proper use of high-quality transformers often requires removal of the direct component from

the primary winding by means of a filter consisting of a choke (or resistance) and blocking condenser (Fig. 9-21). This permits a much better design of transformer but is evidently more costly as well as requiring more space.

Increasing X_p by increasing the number of primary turns has evident limitations, since the secondary turns must be increased also to maintain the turns ratio, and a large increase in secondary turns may so increase C_T as to ruin the h-f response. Thus a compromise must be reached whereby a reasonably large number of primary turns are used with a turn ratio that is not too large, usually between 2 and 4.

Increasing the cross section of the core is quite feasible but tends to make the transformer both bulky and costly. Improvement in core materials to raise the permeability is more desirable, and much

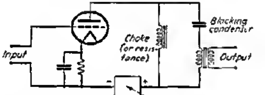


FIG. 9-21. Circuit for removing the direct component of the plate current from the primary of the transformer.

work has been done along this line. Special alloys are now available that have extremely high permeabilities when the transformer rating is not exceeded.

After all means of increasing X_p have been exhausted, any further improvement may be achieved only by a reduction in r_p, which means that the tube selected should not have too high a plate resistance This means further that the amplification factor of the tube cannot be high, since a high-μ tube will also have a high r_p. It is for this reason that triodes with a μ of 8 to 20 are used in this type of amplifier, rather than high-μ triodes or pentodes.

A study of the h-f response, as determined by the h-f multiplying factor of Eq. (9-72), is not so simple, since this fraction, unlike the l-f one, may exceed unity. The study must, therefore, be broken down into two parts: (1) frequencies for which the fraction exceeds unity resulting in the resonant peak shown at about 7500

cycles/sec in Fig. 9-20 and (2) frequencies for which the fraction is less than unity as above 10,000 cycles/sec in Fig. 9-20. For the first part it may be seen that the fraction assumes values in excess of unity because of $(X_T - X_x)$ in the denominator. Therefore, it is desirable that the effect of this term be reduced, which requires high values of r_p and R_x. It will be noted that this conclusion is partly at variance with the requirements for good l-f response which calls for a low value of r_p. An increase in R_x alone avoids this trouble, although it is difficult to insert much resistance into the winding itself. The use of resistance wire in winding the transformer secondary is one method of attempting to improve the h-f response by increasing R_x. Further than this it is possible only to make suitable compromises between the two demands.

The resonant frequency of L_x and C_T should obviously be near or slightly above the maximum frequency to be amplified. This frequency may be increased by a reduction of either or both L_x and C_T; therefore, low leakage reactance is desirable and the shunting capacitance should be kept to a minimum. The distributed capacitance of the transformer constitutes a large part of C_T which may be kept low by suitable winding design, as by winding the secondary in pancakes to maintain a low voltage between adjacent layers of the winding.

It should be evident that the capacitance of C_T in a transformer-coupled amplifier will necessarily be larger than the corresponding capacitance in a resistance-coupled amplifier. For this reason the latter are always preferred for wide-band performance, *e.g.*, in television work where the frequency band may extend to several million cycles per second. Nevertheless transformers may readily be built to provide an amplification which is constant to within 3 db over a range of 40 to 10,000 cycles/sec, and the better class of transformers are capable of extending this range considerably.

It is of interest to note from Eq. (9-72) that the h-f phase-shift term varies from 0 deg at mid-frequency to 180 deg at infinite frequency. Thus the total phase angle of the amplifier varies from 270 deg at zero frequency through 180 deg at mid-frequency on around to 0 deg as the frequency approaches infinity. This is to be contrasted with the resistance-coupled amplifier where the over-all phase angle, as shown by Eq. (9-30), varies from 270 deg at zero frequency around through 180 deg but only to 90 deg as the frequency approaches infinity. This fact is of considerable

importance in feed-back amplifiers, as will be shown in a later section.

Transformers must be shielded to prevent external fields from inducing emfs into their windings and so producing undesired signals or noise in the output. This is especially necessary where the desired voltages are very small, as in the transformers used to couple certain types of microphones to a preamplifier, some of which have magnetic shields up to ½ in. in thickness. Transformers used in the usual run of audio amplifiers do not require so much shielding but should not be mounted too close to a source of intense field, as a 60-cycle transformer supplying power to the plate-voltage rectifier or the tube cathodes.[1]

Resistance Loading in Transformer-coupled Amplifiers. The frequency characteristic of a transformer may be improved at a sacrifice in gain by connecting a suitable resistance across the secondary terminals. The effect of this resistance may be determined by means of the circuits of Fig. 9-19 by placing a resistance R_L, equal to the external resistance divided by the square of the secondary-to-primary turns ratio, across the output terminals of each circuit. Such a resistance will evidently decrease the mid-frequency gain by the ratio $R_L/(r_p + R_s + R_L)$ but the l-f response will be improved since the parallel impedance of $(R_L + R_s)$ and L_p will vary less with frequency than will that of L_p alone. The effect on the h-f response is somewhat more complex, but it is evident that the resonance effect will be reduced by connecting a resistance in parallel with the condenser C_T. The exact performance may be determined, if desired, by developing equations similar to those of Eqs. (9-64) to (9-72).

The curves of Fig. 9-22 show the improvement in the performance of the older transformer of Fig. 9-20 (curve A of Fig. 9-22) when a 100,000-ohm resistance is connected across the secondary terminals (curve B). However, the shunting resistance must not be so low as to drop the gain, in the region of resonance, to a point below the mid-frequency gain.

When a shunting resistance is used, it is frequently in the circuit of Fig. 9-21, where the choke shown in that circuit is replaced by the desired shunting resistance. Thus the resistance not only

[1] For a treatise on shielding, see W. G. Gustafson, Magnetic Shielding of Transformers at Audio Frequencies, *Bell System Tech. J.*, 17, p. 416, July, 1938.

improves the h-f response by reducing resonance, it improves the l-f response both by producing the effect of a lower r_p and by increasing L_p through a reduction in the saturation of the iron core.

Impedance-coupled Amplifiers. Voltage amplifiers may also be built by replacing R_1 in the circuit of Fig. 9-4 by an inductance. Such amplifiers have the advantage over resistance-coupled amplifiers of requiring a lower direct-voltage source to secure the desired plate voltage but have the disadvantage of a poorer frequency response due to the variations of the impedance of the coupling unit with frequency. The coupling impedance must be high enough at the lowest frequency to be large compared to r_p, just as was required of the reactance of L_p in the transformer, while

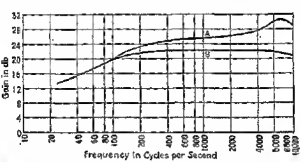


FIG. 9-22. Effect of resistance loading on the frequency characteristic of a transformer-coupled amplifier. *A*, no resistance across secondary; *B*, 100,000 ohms across secondary.

the distributed capacitance must be small enough to provide satisfactory h-f response. In many respects the performance of such an amplifier is essentially that of a transformer-coupled amplifier with a 1:1 turns ratio. Impedance-coupled amplifiers are subject to most of the disadvantages of transformer coupling without the advantages of higher gain due to the step-up action of the transformer and are therefore not widely used.

Cathode-ray Oscilloscopes.[1] The cathode-ray tube described

[1] This subject is introduced at this point because the oscilloscope is a tool with which anyone studying vacuum-tube amplifiers should be familiar. If laboratory work accompanies the study of this text, the oscilloscope will serve as an excellent piece of test equipment for studying the circuits described in this and succeeding chapters.

in Chap 6 (page 145) may be combined with vacuum-tube amplifiers and a linear sweep circuit to provide an effective measuring and analyzing unit. Such a combination is commonly known as an *oscilloscope*, the block diagram of Fig. 9-23 indicating the general arrangement. Amplifiers, inserted between the input terminals to the oscilloscope and the "horizontal deflection" and "vertical deflection" terminals of the circuit shown in Fig. 6-3 (page 148), enable a very small signal voltage to produce a large movement of the spot on the screen of the tube. The sweep oscillator, when used, is normally applied to the horizontal deflection plates through the associated amplifier and "sweeps" the spot across the screen horizontally at a uniform rate. The wave shape of such an oscillator should appear as in Fig. 9-24 to cause the spot to move uniformly from left to right on the screen

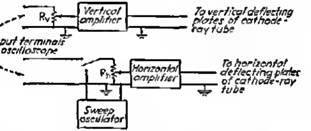


FIG 9-23 General arrangement of the component parts of an oscilloscope

from time a to b and return to the left side of the screen instantly from b to a' In practical oscillators the section ba' is not perfectly vertical but is nearly so if the frequency is not too much above the a-f spectrum.

A simple type of sweep circuit uses a cold-cathode tube as a relaxation oscillator. In its basic form such an oscillator consists of a condenser, resistance, gas-filled tube (*e.g.*, such as one of the cold-cathode tubes described on pages 117 to 119) and a source of potential (Fig. 9-25). When the battery circuit is closed, the condenser C charges at a rate depending upon its capacity and the size of the resistor R As soon as the voltage across the condenser becomes sufficiently high, tube T breaks down and virtually short-circuits the condenser, dropping its potential to the point where conduction through the tube ceases. The con-

denser then again builds up slowly, and the cycle is repeated. The condenser voltage, therefore, follows a saw-tooth curve as in the heavy line of Fig. 9-26, where a and c represent the potential at which the tube ceases to conduct and b is its breakdown potential. The dotted line shows the curve that the condenser voltage would follow if no tube were used, assuming that the battery potential is very much higher than the breakdown point of the tube. A

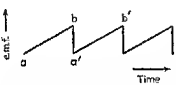


FIG. 9-24

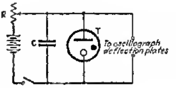


FIG. 9-25

FIG. 9-24. Ideal wave shape of a sweep-circuit oscillator.
FIG. 9-25. Sweep-circuit oscillator using a cold-cathode tube. The frequency is varied by changing either R or C.

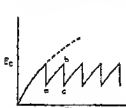


FIG. 9-26

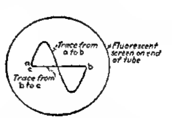


FIG. 9-27

FIG. 9-26. Curve of condenser voltage vs. time for the circuit of Fig. 9-25.
FIG. 9-27. Sine wave as it appears on the screen of a cathode-ray tube equipped with a linear sweep circuit. Note the return trace. It is not always at the zero line as shown.

high battery potential is desirable, as it tends to make the portion of the curve ab more nearly a straight line.

When used with a cathode-ray tube, the frequency of the saw teeth should be an exact multiple of that of the current, or voltage, to be observed, or the image will not be stationary. Figure 9-27 shows the appearance of the screen of a cathode-ray tube with a sine wave impressed across the vertical deflecting plates and a sweep-circuit oscillator across the horizontal plates, when the frequency of the latter source is exactly equal to that of the

sine-wave source. The points a, b, and c represent the same points of the saw-tooth cycle as in Fig. 9-26. The straight line from b to c is made by the rapid change in potential when the glow tube breaks down and is not a zero line. By shifting the phase of the timing wave with respect to the observed wave this line may be shifted up or down at will. As a matter of fact, this line is usually so faint as to be barely discernible.

The desirability of maintaining synchronism between the sweep and observed circuits has led to the use of small thyratrons as the discharge tube. These thyratrons are generally filled with argon gas, rather than mercury vapor, ensuring a breakdown potential that is independent of the temperature. They are commonly used in a circuit similar to that of Fig. 9-28. Synchronism is obtained by exciting the grid of the thyratron from the same source

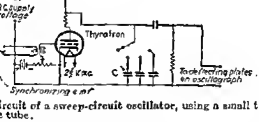


Fig. 9-28. Circuit of a sweep-circuit oscillator, using a small thyratron as the discharge tube.

as is applied to the vertical deflecting plates. The grid prevents the tube from conducting until the grid potential is sufficiently positive, so that if the R and C of the discharge circuit are adjusted to near synchronism, the grid will trip the tube off at even intervals of one cycle. It is also possible to adjust R and C to give a frequency of approximately $1/n$ times the impressed frequency (where n is any whole number) and so provide synchronism at subharmonics. This will show more than one cycle of the impressed wave, the number of cycles being equal to n.

Cathode-ray oscilloscopes may be used to observe the voltage wave at any point in a vacuum-tube amplifier and thus permit a quick check on the performance of the amplifier without disturbing its operation, since the input resistances of the oscilloscope (R_a and R_b, Fig. 9-23) are of the order of at least hundreds of thousands

of ohms and may be bridged across the plate-load impedance of most amplifier tubes.[1]

It is often desirable to apply another test voltage to the horizontal pair of deflecting plates of a cathode-ray tube instead of a sweep-circuit oscillator (by throwing the switch in Fig. 9-23 to the upper position). The resulting Lissajous figures, when properly interpreted, often provide valuable information which cannot be obtained in any other way.[2]

D-C Amplifiers. None of the circuits so far discussed is suitable for amplifying direct voltages or alternating voltages of a frequency less than a few cycles per second. The simplest circuit for such service is obtained by modifying the ordinary resistance-coupled amplifier as shown in Fig. 9-29. The battery E_{cc_2} is used to reduce the voltage on the grid of the second tube to its required negative potential.

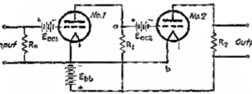


FIG. 9-29. Simple type of d-c amplifier.

In operation any change of potential in a positive direction at the input will cause an increase in plate current in tube 1. This will produce a greater drop through R_1 and therefore a more negative grid voltage and a lower plate current in tube 2. If more stages are added, each succeeding tube will be found to work alternately; *i.e.*, the plate current will increase in the third tube as in the first, decrease in the fourth tube as in the second, etc.

[1] If the signal is sufficiently strong, the deflecting plates of the cathode-ray tube may be bridged directly across the amplifier under test, whence the bridging impedance is merely that of the small capacitance between deflecting plates.

[2] A study of the figures that may be obtained on an oscilloscope with different combinations of applied voltages (*e.g.*, a sine wave on one pair and its second harmonic on the other) is beyond the scope of this book. An excellent treatise on this subject is given by John F. Rider, "The Cathode-ray Tube at Work," John F. Rider, New York.

Each tube will, of course, produce a greater change in plate current than its predecessor, owing to its amplification.

A circuit that avoids the use of the extra grid batteries, is shown in Fig. 9-30.[1] The grid bias on tube 1 is the drop between A_1 and C_1. That on tube 2, however, is equal to the difference between the negative voltage developed across the resistance R_1 by the plate current of the first tube and the positive potential between points A_2 and B_1. The bias on tube 3 is obtained in the same manner as that for tube 2. The anode voltage for each tube is equal to A_1B_1, A_2B_2, or A_3B_3 (depending on the tube) less the I_bR drop through the coupling resistance.

The tube heaters may each be energized from a separate source, or a common source may be used if the insulation between each cathode and its heater is sufficient. The resistor $C_1A_1 \ldots B_3$ is

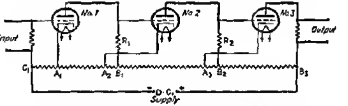


FIG. 9-30. Loftin-White type of d-c amplifier.

used as a voltage divider and must be capable of passing sufficient current to stabilize the circuit. The direct voltage supplied must be equal to

$$\text{D-C supply} = E_{c_1} + E_{b_1} + E_{c_2} + E_{b_2} + E_{c_3} + E_{b_3} + I_{b_3}R_3$$

Note that it is not equal to $E_{c1} + E_{bb1} + \cdots$, since it is not necessary to provide additional emf for the d-c drop through the load resistors, except for the last tube.

All d-c amplifiers are subject to "drift"; i.e., a change in the supply potential for any tube will cause the currents and potentials of all succeeding tubes to vary. If the grid potential of the first tube, for example, varies slightly, the gain of the amplifier will cause the current in the last tube to vary a large amount, even to

[1] Edward H. Loftin and S. Young White, Direct-coupled Detector and Amplifier with Automatic Grid Bias, *Proc. IRE*, 16, p. 281, March, 1928.

the point of decreasing to zero and so making the amplifier inoperative, or increasing to an excessively high value. In circuits like that of Fig. 9-30 drifting is minimized by the use of a voltage-regulated power supply and by using a bleeder resistance which is as low as can be economically used.

Drifting may also be reduced by using a balanced amplifier as in Fig. 9-31. Here the output voltage is the result of an unbalance between two tubes operating in push-pull. If the two tubes in each pair have exactly similar characteristics and the circuit constants are exactly matched, this unbalance can result only from a voltage applied across the input terminals. Any

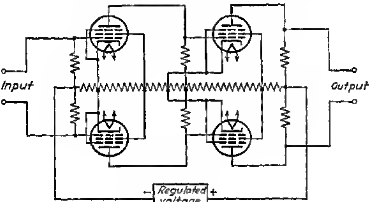


FIG. 9-31. Balanced type of d-c amplifier.

change in the direct voltage supply to the grids and plates of the tubes will have equal effects on the two tubes of a pair and will produce no unbalance. In practice perfect balance is not possible but, when a regulated power supply is used with a balanced amplifier, remarkably stable results are obtained.[1]

While triode tubes are shown in the circuit of Fig. 9-30 and pen-

[1] For a more detailed discussion, see such references as Franklin F. Offner, Balanced Amplifiers, *Proc. IRE*, **35**, pp. 306–316, March, 1947; Harold Goldberg, Bioelectric-research Apparatus, *Proc. IRE*, **32**, pp. 330–334, June, 1944; Harold Goldberg, A High Gain D-C Amplifier for Bioelectric Recording, *Trans. AIEE*, **59**, pp. 60–64, January, 1940.

todes in Fig. 9-31, either type may be used in either circuit. Pentodes are usually preferred because of their higher gain.

Amplifiers Controlled by Phototubes. It is frequently desirable to amplify the output of a photoelectric tube. Any of the amplifiers described in this chapter may be used for this purpose if a high resistance is inserted in series with the phototube to produce the voltage necessary to drive the amplifier.

The frequency response of an amplifier to be used with a phototube must be such as to handle the rate of change of light that may be expected. Thus if the light on the phototube is to be flashed on and off at a very low rate, such as once or twice per second, a d-c type of amplifier must be used. On the other hand if the light is to vary over a wide frequency range, as in the reproduction

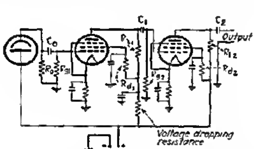


FIG. 9-32. Resistance-coupled amplifier circuit for reproducing the output from a phototube, as in the reproduction of sound-on-film recordings.

of sound-on-film recordings, the frequency response of the amplifier must be reasonably uniform over the entire audible spectrum covered by the recording. The circuit of Fig. 9-32 illustrates the use of a resistance-coupled amplifier to amplify the output of a phototube picking up light from a sound-on-film recording (or other similar source). Resistance R_0, in series with the phototube, sets up the voltage to be amplified by the first tube. The condenser C_0 removes the direct component produced by the average light intensity impressed on the phototube, so the voltage impressed on the grid of the first tube represents *variations* in the light only rather than the absolute value.

The resistance R_0 must be very large, of the order of at least 5 or 10 megohms and preferably more, to produce the highest possible

voltage on the grid of the first tube. Since R_{g1} and the grid-cathode circuit of the first tube are effectively in parallel with R_0 as far as any alternating components are concerned, the resistance of each must be at least as high as and preferably higher than R_0; therefore, the grid current in the first tube must be extremely small. Since the grid is maintained negative, any grid current that may flow must be due to the presence of positive ions, so only a low plate voltage (and also screen voltage in pentodes) should be applied to this tube to prevent ionization of any traces of gas present, even at a sacrifice in gain. This is especially true if a high-vacuum phototube is used, as its lower sensitivity and higher internal resistance make a higher value of R_0 necessary than for the gas-filled tube if equal response is to be realized. This is a major reason for the more general use of gas-filled tubes in sound reproduction than of high-vacuum tubes.

The same type of amplifier may be used even when the light on the phototube varies at a frequency of the order of a few cycles or less, by interrupting the light beam at a much higher rate. Thus the frequency at which the light is interrupted represents the carrier of a modulated wave (see Chap. 13 for details of modulation) and is sufficiently high to be readily amplified by the circuit of Fig. 9-32; the l-f variations of light intensity represent the modulation carried on the carrier. By operating the final tube of the amplifier on the curved portion of its curve, to serve as a demodulator, it is possible to recover the original l-f signal in the output. This avoids the use of a d-c amplifier with its attendant problems of drift and instability.[1]

Volume-control Methods. The output of resistance-coupled amplifiers is most conveniently varied by replacing the grid leak in one stage with a high-resistance potentiometer (Fig. 9-33). This has very little effect on the frequency response of the amplifier and permits control of the output signal from zero to maximum level.

The same arrangement is shown applied to a transformer-coupled amplifier in Fig. 9-34. The potentiometer has the same effect upon the performance of the transformer as does resistance

[1] This amplifier is essentially another application of the principles presented by L. J. Black and H. J. Scott, A Direct-current and Audio-frequency Amplifier, *Proc. IRE*, **28**, p. 269, June, 1940.

loading (see page 302) and may actually improve the frequency characteristic while somewhat decreasing the maximum gain.

Volume-limiting, volume-compressing, or volume-expanding characteristics may be added to an audio amplifier. As an example, sudden overloads on amplifiers used in public-address systems and in radio-broadcast studios, as when a speaker suddenly raises his voice or places his mouth too close to the microphone, may be avoided by using a "chopping" device which limits the

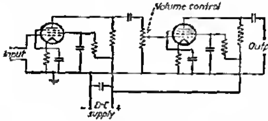


FIG 9-33. One method of controlling the gain of a resistance-coupled amplifier

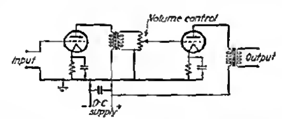


FIG 9-34 One method of controlling the gain of a transformer-coupled amplifier.

maximum output of an amplifier no matter how large the impressed signal may be

A more elaborate arrangement is to design a variable-gain amplifier, in which the gain decreases as the input signal rises.[1] Thus the output will contain all the original input components,

[1] This may be done, for example, by rectifying a portion of the output signal to secure a direct component that is proportional to the average output taken over a fraction of a second. This direct component is then used to control the bias of remote-cutoff pentodes used in one or more stages of resistance-coupled amplification, in a manner similar to the automatic-volume-control circuits used with r-f amplifiers (p. 587).

but the volume range will be compressed. This is particularly advantageous where the signal is to be transmitted any appreciable distance over telephone lines. As the signal on a telephone line becomes weak, the possibility of cross talk from other paralleling lines increases; if the signal becomes too strong, the circuit may cause cross talk into adjacent channels. Broadcast stations use considerable care to maintain the energy level within rather close limits on the telephone lines carrying their programs; and unless an automatic compressor is used, this must be done by hand. Another application in which volume compression is highly desirable is in recording where too strong a signal may cause the recording stylus (in the case of disk recordings) to cut over into the adjoining groove or may cause overloading of the light valve in sound-on-film recording.

Volume expanders should be used when receiving a signal to which a volume compressor has been applied. An amplifier so equipped will have a variable gain in which the gain increases with the intensity of the impressed signal. If the characteristics of the expander are complementary to those of the compressor, the resulting output should be identical in volume range with the signal originally impressed on the input of the compressing amplifier. Unfortunately all compressors and expanders require some form of time-delay circuit to prevent them from following each individual cycle; therefore, the output of the expander may not follow rapid changes in volume of the original signal. Nevertheless, quite satisfactory results are obtainable in most commercial applications.

Circuits for volume compressors and expanders are rather numerous and, in many cases, complex. The reader is, therefore, referred to the technical literature for detailed circuits.[1]

[1] The following are samples of the literature on this subject: R. C. Mathes and S. B. Wright, The Compandor—An Aid Against Radio Static, *Elec. Eng.*, **53**, p. 860, June, 1934; S. B. Wright, Amplitude Range Control, *Bell System Tech. J.*, **17**, p. 520, October, 1938; S. B. Wright, S. Doba, and A. C. Dickieson, A Vogad for Radiotelephone Circuits, *Proc. IRE*, **27**, p. 254, April, 1939; E. G. Cook, A Low-distortion Limiting Amplifier, *Electronics*, **12**, p. 38, June, 1939; H. H. Stewart and H. S. Pollock, Compression with Feedback, *Electronics*, **13**, p. 19, February, 1940; C. G. McProud, Volume Expansion with a Triode, *Electronics*, **13**, p. 17, August, 1940; and H. H. Scott, Dynamic Noise Suppression, *Electronics*, **20**, p. 96, December, 1947.

Sources of Noise in Amplifiers. In general, the only limit to the amplification that may be employed in any given application is imposed by the signal-to-noise ratio. Any noise will be amplified along with the desired signal so that if the noise level at any point in the amplifier is an appreciable proportion of the useful signal, further amplification will be worthless. It is, therefore, particularly important that the noise level be kept low in the early stages of a high-gain amplifier.

Noise may be introduced by poor contacts, worn-out batteries, carbon resistors, or any other source of irregularity of current flow, but even when such sources are quite carefully eliminated, a limit is imposed on the maximum useful gain of a high-gain amplifier by *thermal agitation*. Current flow actually consists of the movement of discrete particles, electrons. These are known to move around in a somewhat irregular manner within the confines of the conductor even when there is a gradual drift (or current flow) in a given direction, the magnitude of this movement being a function of the temperature. Such electronic movements set up small voltages which are entirely random in frequency, so the energy impressed on the grid of an amplifier is spread quite uniformly over a very wide frequency band, much wider than the presently used r-f spectrum. Thus thermal agitation imposes a limit on the maximum gain of any amplifier, whether r-f or a-f.

Another source of noise is the *shot effect*. Since an electric current is actually a flow of electrons, the plate current in a tube consists of a series of small particles impinging on the plate rather than a smooth flow as normally considered. In most amplifiers the magnitude of the alternating current flowing, even in the early stages, is sufficient to mask the effect of individual electrons, but in very high-gain amplifiers the signal applied to the input circuit may be so low that the small irregularities in plate current, owing to the individual electrons striking the plate, will produce an appreciable noise voltage in the output of the amplifier. It has been found that space charge tends to smooth out this effect, and a considerable excess of emission over and above the normal flow of plate current is, therefore, desirable in any tube used in the early stages of a high-gain amplifier.

Evidently neither thermal agitation nor the shot effect is of any importance unless the desired signal is very small, since only then will the signal-to-noise ratio be low. Thus their principal

effect is to impose limits on the maximum useful amplification obtainable.

Hum. Hum is the name given to induction in an amplifier from the 60-cycle supply and its harmonics. Inadequate filtering of the rectifier that supplies the direct voltages is a common source of hum, the frequency of which is usually 120 cycles/sec due to the use of single-phase, full-wave rectifiers.

Hum may also be introduced by the alternating current used to heat the cathodes. Filamentary-type cathodes, if heated from an a-c source, must be supplied with a mid-tap, usually a mid-tap of the transformer secondary supplying the heater current, to which the grid and plate return leads are connected. Since the average potential of this mid-tap remains unchanged with respect to the various parts of the filament, hum from this source is reduced to a minimum. Most cathodes are, of course, indirectly heated because of the much lower hum level possible with this type. Some hum may be present even so, owing to induction either from the leads and other current-carrying parts or from potentials set up between the cathode and heater. It is essential that the heater be tied to the cathode, preferably by means of a mid-tap on the transformer secondary supplying the heater current, to prevent alternating voltages being set up between cathode and heater. Under some conditions it may even be desirable to make the heater negative with respect to the cathode. A twisted pair should always be used for the wiring between the cathode heaters and the transformer, to reduce induction from the alternating heater current.

A more common source of hum is electromagnetic induction into various parts of the amplifier, especially into the transformers of a transformer-coupled amplifier. The use of resistance coupling in the early stages of a high-gain amplifier will largely eliminate this trouble, but a transformer must frequently be used in the input circuit to match the high impedance of the amplifier to the lower impedance of a transmission line or microphone. Such a transformer must be very heavily shielded to eliminate all hum, since if even a very small amount of hum is introduced at this point, the signal-to-noise ratio will be low in all the remaining stages of amplification. It is also possible to reduce the amount of hum picked up in this and other transformers by so arranging them with respect to the sources of a-c induction, such as power transformers

and chokes in the power-supply unit, as to minimize the pickup. An experimental method of determining the proper location is often the most satisfactory.

Electrostatic induction is also a source of hum. Any point in the amplifier that is separated from ground by a high impedance is subject to electrostatic induction, since even a small current will then set up a comparatively high voltage. The grid circuits of the various tubes in the amplifier are most susceptible to such induction, since they are usually separated from ground by high-impedance grid leaks (the cathode normally being at or near ground potential). The grid leads of the earlier stages of a high-gain amplifier must therefore be short or thoroughly shielded to prevent hum pickup.

Microphonic Noise. Microphonic noise is caused by vibration of the tube elements, owing either to mechanical vibration transmitted through the tube socket or to sound waves striking the tube envelope. Vibration of a tube due to either cause produces variations in the spacing between the electrodes and, therefore, in the space currents. These variations, being at audible frequencies, produce a sound in the loud-speaker.

Microphonic noise may be reduced by proper design of the tube and of its mounting. Some tubes have less microphonic action than others, certain types being constructed especially for the early stages of high-gain amplifiers where the effects of microphonic action are most serious. Mounting of the entire amplifier on rubber cushions and, if necessary, separate mounting of the tube sockets will reduce the probability of microphonics.

It should be here noted that although the noise and hum discussed in the preceding sections have been largely of an a-f nature, they are likely to cause considerable trouble in r-f amplifiers as well. All vacuum tubes have some nonlinearity in their characteristic curves which, as shown in Chap. 13, will cause the disturbing signal to be modulated on the r-f currents and so pass on through the amplifier in the same manner as the desired a-f modulation carried by the r-f currents.

2. POWER AMPLIFIERS

Nearly every amplifier tube produces a power output that is greater than the power delivered to its grid and is, therefore, in the strict sense of the word, a *power amplifier*; but this term has

been generally applied to those tubes used to supply power to circuits wherein power, rather than voltage, is the principal consideration. Tubes supplying loud-speakers, radio antennas, etc., are examples of this class of amplifiers.

Power amplifiers may be classed as r-f or a-f. The former are usually either class B or class C and are described in the next chapter. The latter are either class A or, when operated in push-pull, class B or class AB. The performance of a-f amplifiers is covered in this part of the present chapter.

In studying power amplifiers considerable emphasis will be placed on methods of computing and reducing amplitude distortion. Very little was said about amplitude distortion in the preceding sections on voltage amplifiers because the amplitude of the signal in a voltage amplifier is likely to be very much less than the maximum capacity of the tube so that amplitude distortion, when the tube is operated with correct electrode voltages, is usually negligible in magnitude. The final stage of a voltage amplifier, however, may normally be expected to produce a voltage close to the maximum capacity of the tube, and in such cases the methods of the succeeding sections may be used to determine the amplitude distortion to be expected. Similarly the equations and procedures previously given for determining the frequency distortion in voltage amplifiers may be applied to power amplifiers and some discussion of their application will be given later (page 357).

Class A Power Amplifiers. The study of class A power amplifiers involves the determination of their power output, the distortion present, and particularly the optimum load resistance and direct electrode potentials for maximum power output without excessive distortion. If a set of static characteristic curves of the tube is plotted together with the dynamic curve for a given load resistance, the performance of the tube may be determined. Such a set of curves was shown in Fig. 3-21, page 57, and is reproduced in Fig. 9-35 for convenience. Here it is evident that, with a sine wave of emf impressed on the grid, the plate-current wave shape is definitely not sinusoidal; *i.e.*, amplitude distortion is present (see page 258 for definitions of the various types of distortion). It is further apparent that the cause of this distortion is the curvature of the characteristic curves, particularly at low values of plate current. Therefore, one requirement for distortionless amplification is that the plate current must never, at any part

of the a-c cycle, be permitted to swing down into the appreciably curved portions of the characteristic curves.

Distortion may also appear if the grid is permitted to swing positive at any point in the cycle. Grid current flows, indicating a decrease in the input resistance of the tube, whenever the grid swings positive. If the driving amplifier is resistance coupled,

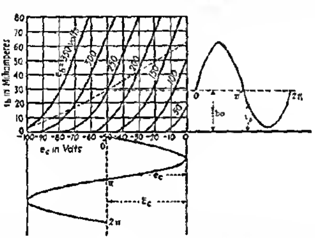


FIG. 9-35. Use of the dynamic curve to determine the performance of an amplifier.

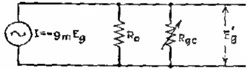


FIG. 9-36. Equivalent circuit of a driving amplifier illustrating the effect of grid current flowing in the driven tube.

this is equivalent to lowering the grid-leak resistance R_g, Fig. 9-4, during a portion of the cycle which, as shown by Eq. (9-30), will decrease the gain of the driving amplifier during that period of time by decreasing R_0. The same analysis may be made of a transformer-coupled amplifier leading to similar conclusions

The problem is more simply put in Fig. 9-36, which represents the equivalent circuit of a resistance-coupled driving amplifier in the

mid-frequency range. The variable resistance R_{gc} represents the grid-to-cathode resistance of the driven tube, which is high throughout that part of the a-c cycle that the grid is negative but which decreases materially during a small portion of the cycle when the grid goes positive.[1] It is evident that the voltage E_g', which is applied to the grid of the driven tube, will be equal to $-g_mE_g\left(\frac{R_0R_{gc}}{R_0+R_{gc}}\right)$, showing that if R_{gc} remained high through- the cycle, E_g' would be equal to $-g_mE_gR_0$ as shown by the solid curve of Fig. 9-37. But when the grid swings positive, R_{gc} becomes smaller than R_0 causing E_g' to follow the dashed curve.

There are several solutions to the foregoing problem, the most common of which, for class A amplifiers, is to maintain the grid negative at all times by applying a suitable d-c bias. The input resistance then remains high at all times throughout the cycle,

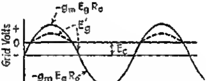


FIG. 9-37. Illustrating the distortion introduced by permitting grid current to flow in the driven tube.

and no distortion results. When the grid is permitted to swing positive, other methods must be employed to reduce distortion. Such cases will be treated in a later section on class B amplifiers (page 350), as class A amplifiers are seldom so operated.

The foregoing may be summarized by stating that distortionless operation of a class A power amplifier normally requires that the grid be maintained negative and that the plate current be kept above a given minimum throughout the cycle of impressed voltage.

Conditions for Maximum Power Output. Resistance Load. The quantitative study of a power amplifier may be simplified by using the plate-current–*plate*-voltage characteristics rather than the plate-current–*grid*-voltage characteristics of Fig. 9-35. Such a set for a triode is shown in Fig. 9-38. The operation of the amplifier, as summarized in the preceding paragraph, must be

[1] In developing Eq. (9-30) R_{gc} was considered infinite.

confined to the shaded area, the upper edge of which is the zero grid voltage curve, whereas the lower edge represents the minimum permissible plate current. This minimum current is a function of the permissible distortion and of the resistance of the load. It is therefore not a rigidly fixed quantity but must be determined for the conditions under which the amplifier is to work. Nevertheless its range of variation is not great, and the use of an arbitrarily fixed minimum is of great assistance in visualizing the performance of the amplifier.

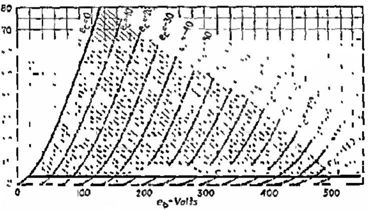


FIG. 9-38. Distortionless operation of a class A amplifier requires that variations in plate current and voltage be restricted to the crosshatched area.

Let it be assumed that the tube is to operate in the circuit of Fig. 9-39, that T_2 is an ideal transformer at the frequencies to be used which presents an impedance across its primary terminals of $R_L = (N_1/N_2)^2R_s$ (where N_1 and N_2 are the number of primary and secondary turns, respectively, of the transformer), and that the impedance of C_c is zero. Let it be further assumed that the grid bias (voltage drop across R_c) is 50 volts and that the direct plate voltage (d-c supply less the drop across R_c, the resistance drop in the transformer primary being negligible) is 250 volts. If the alternating voltage supplied to the primary of the grid-excitation transformer T_1 is, for the moment, zero, only direct current will flow in the plate circuit, the magnitude of which may

be determined from the static curves (replotted in Fig. 9-40) by the intersection of the −50-volt grid curve with the 250-volt plate-voltage line, an intersection that gives $I_b = 29$ ma.

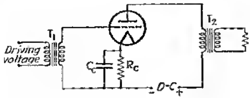


FIG. 9-39. Basic circuit of a class A power amplifier.

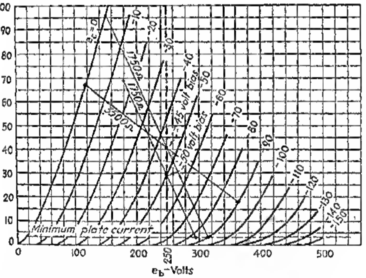


FIG. 9-40. Illustrating the effect of improper bias on the performance of a class A amplifier.

When alternating voltage is applied to the primary of the excitation transformer T_1, the plate current and plate voltage will vary above and below their direct values of 29 ma and 250 volts, respectively, the plate voltage increasing as the current decreases. We may determine the relationship, and so plot the curve of i_b

vs e_b, by writing the following equations taken from pages 63 and 64 and the footnote on page 67

$$i_b = I_b - i_p \tag{9-73}$$

$$e_b = E_b + e_p \tag{9-74}$$

$$e_p = i_p R_L \tag{9-75}$$

By combining these three equations we may write the equation of the i_b-e_b curve, known as a *load line*,[1]

$$e_b = E_b + I_b R_L - i_b R_L \tag{9-76}$$

When corresponding values of e_b and i_b, as obtained from this equation for $R_L = 1750$, $E_b = 250$, $I_b = 0.029$, are plotted in Fig. 9-40, the lower of the two 1750-ohm lines is obtained. This load line is always straight, when a resistance load is used, as contrasted with the curved nature of the dynamic curve of Fig. 9-35.

It should be noted that the lower 1750-ohm load line of Fig. 9-40 is extended to, but not below, the minimum plate current (compare with the minimum plate-current line of Fig. 9-38) which permits a maximum negative grid swing of about 35 volts (from -50 to -85). Obviously the maximum swing in the positive direction with a sinusoidal input is also 35 volts, and the curve is therefore extended in a positive direction to -15 volts on the grid. To secure the maximum possible output without excessive distortion the alternating voltage should be such as to decrease the plate current just to the minimum value on the negative half cycle and yet reach zero grid voltage on the positive half cycle. The grid bias of -50 volts must, therefore, be slightly reduced until the load line intersects the minimum plate-current line at a grid voltage of twice the bias. This is illustrated by the upper 1750-ohm load line in Fig. 9-40 where the bias has been reduced to -45 volts permitting an excitation voltage of 45 volts crest value without excessive distortion. The grid swing is from 0 to -90; a voltage of -90, or twice the bias, reducing the current to a point just slightly into the curved region.

The procedure just outlined may be followed for other loads by

[1] Since the relationship between i_b and e_b is determined entirely by the load impedance and not at all by the tube characteristics, this curve is known as a *load line* rather than a dynamic curve (the term applied to the dotted curve of Fig. 9-35).

so adjusting the excitation and bias as to maintain constant distortion (the crest value of the excitotion voltage being at all times equal to the bias). The power output and bias so obtained are plotted against load resistance in Fig. 9-41.[1]

Theoretical Determination of Load Resistance to Be Used.[2] It is possible to compute the load resistance which will produce the maximum power output under the assumption of straight-line characteristic curves. Such an analysis will not be exactly correct for actual amplifiers because of the curvature of the characteristic curves of the tube; nevertheless the information obtainable in this manner is useful and the results do not differ too much from those

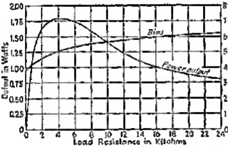


FIG. 9-41. Curves of output and bias of a triode tube for constant 5 per cent distortion. The crest value of the signal voltage is equal to the bias. (r_p = 1600 ohms)

of less general but more accurate methods such as the graphical analysis starting on page 328.

Figure 9-42 shows a set of hypothetical tube static characteristic curves which are perfectly straight. A load line CF is drawn in a manner similar to that employed in drawing the load lines in Fig. 9-40. The distance OB represents the plate voltage E_b which is assumed to be held constant. For maximum output without distortion the grid voltage should swing just to zero on the positive

[1] Methods of computing the power output and distortion are outlined in a later section beginning on p. 328. The 3500-ohm load line shown in Fig. 9-40 has nothing to do with the present discussion but is referred to on page 328.

[2] For additional discussion of this problem, see Wayne B. Nottingham, Optimum Conditions for Maximum Power in Class A Amplifiers, *Proc. IRE*, 29, pp. 620-623, December, 1941.

half cycle of the excitation voltage and sufficiently negative on the negative half cycle to carry the plate current just to the curved region. Since it was assumed that the plate current curves are straight all the way to zero, the negative swing of the grid voltage should be just sufficient to cause the plate current to drop to zero. This means that operation is confined to the region CDE of the load line, where D is so located that $CD = DE$, and we may write $I_b = HO$, $I_{pm} = HO = GH$, $E_b = OB$, $E_{pm} = AB = BC$.

Since the objective of this development is to determine the value of the load resistance, R_L, which will give the maximum power

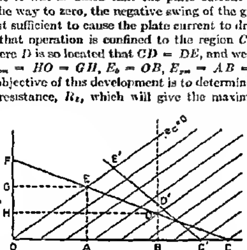


FIG. 9-42. Theoretical straight-line characteristic curves. Minimum plate current equal to zero.

output without distortion, the first step is to write the equation for the power output

$$P = I_p^2 R_L = \frac{I_{pm}^2 R_L}{2} \tag{9-77}$$

Maximum power output is found by differentiating P with respect to R_L and equating the result to zero. Since the current I_{pm} will vary with R_L, it must be expressed in terms of R_L before we can differentiate. To do so let us first write

$$I_{pm} = \frac{E_{pm}}{R_L} \tag{9-78}$$

E_{pm} in the above equation may be expressed in terms of E_b by writing

$$E_{pm} = E_b - OA \tag{9-79}$$

The distance OA may be expressed in terms of AE and r_p (since r_p is the reciprocal of the slope of the line OE) and AE is evidently equal to $2I_{pm}$. Therefore

$$r_p = \frac{OA}{AE} = \frac{OA}{2I_{pm}} \tag{9-80}$$

Solving this equation for OA and substituting in Eq. (9-79) gives

$$E_{pm} = E_b - 2I_{pm}r_p \tag{9-81}$$

Substituting Eq. (9-81) into Eq. (9-78) and solving the resulting equation for I_{pm} gives

$$I_{pm} = \frac{E_b}{R_L + 2r_p} \tag{9-82}$$

When Eq. (9-82) is substituted into Eq. (9-77) the result is

$$P = \frac{E_b^2 R_L}{2(R_L + 2r_p)^2} \tag{9-83}$$

The only variable on the right-hand side of this equation is R_L so we may differentiate P with respect to R_L and set the result equal to zero to determine the conditions for maximum power output. When this is done, we find that

$$R_L = 2r_p \tag{9-84*}$$

Thus, under idealized conditions of straight-line characteristic curves the load resistance should be made equal to twice the plate resistance of the tube to obtain maximum power output without distortion.

It may seem strange that the theoretical load resistance for maximum power output is twice the internal resistance of the tube instead of being equal to it. For example, let us rewrite Eq. (9-77) by first solving Eq. (3-20) (page 63) for I_p and putting this value of I_p into Eq. (9-77). This will give

$$P = (\mu E_g)^2 \frac{R_L}{(R_L + r_p)^2} \tag{9-85}$$

and when P is differentiated with respect to R_L and the result equated to zero, we find that maximum power output occurs when $R_L = r_p$. The difference in these two results is that Eq. (9-83) was developed for maximum output *without distortion* assuming

the necessary excitation voltage to be available. It is therefore based on the assumption that the minimum plate current is equal to zero and that operation on the load line takes place between two *nonparallel* lines, OE and OC, the bias (direct grid voltage) and excitation (alternating grid voltage) being suitably readjusted with each change in R_L so as to realize this condition.

This is further illustrated by the second load line $C'D'E'$ which represents a lower load resistance. The operating point has been shifted from D to D' by a change in bias, and the excitation voltage has also been changed so that its crest value is equal to the bias, $E_{gm} = E_c$. The point D' is so located that $C'D'$ equals $D'E'$, thus operation over the load line is again confined to the region between the lines OE and OC.

Equation (9·85), on the other hand, was developed for maximum output for a *fixed excitation voltage* and is based on the assumption that the characteristic curves are both straight and *continuous* throughout the region of plate-current variation, so that distortion is not a problem. The requirement of continuity is met by making the minimum plate current greater than zero, as illustrated in Fig. 9-43, where the bias is such as to set the operating point at D. If the crest value of the excitation voltage is again made equal to the bias, operation will be between points E and N and the minimum plate current will be OJ. A lower value of load resistance is represented by the line $E'DN'$, no change in bias or excitation voltage being made. It may be seen that operation under these conditions is between two *parallel* lines, OE and PN, and under these conditions maximum power output requires that $R_L = r_p$.

These two conditions are further illustrated by the power output curve of Fig. 9-41, computed for an actual triode tube for a constant distortion of 5 per cent and thus corresponding to the construction of Fig. 9-42, and the power output curve of Fig. 9-44, taken for the condition of constant excitation voltage and thus corresponding to the construction of Fig. 9-43. (In the latter case the crest value of the excitation voltage was less than the bias to keep point N' above the curved region of the static curves so as to prevent distortion. This has the effect of lowering the output but is necessary to prevent distortion from changing the shape of the output curve.) Comparison of these two curves shows that the maximum power output in Fig. 9-41 occurs at approximately

$R_L = 2r_p$ (r_p = 1600 ohms) and that maximum power output in Fig. 9-41 occurs at $R_L = r_p$.

The statement was made in a preceding paragraph that the minimum plate current in the construction of Fig. 9-42 was zero. This is of course true only in the idealized case of perfectly straight

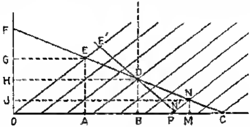


FIG. 9-43. Operation on the theoretical curves of Fig. 9-42 when the minimum plate current is greater than zero.

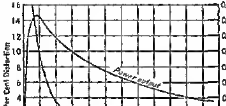


FIG. 9-44. Curves of output and distortion for a triode tube for constant impressed signal voltage; E_b = 250 volts, E_c = −50 volts, r_p = 1600 ohms. (Only 20 volts crest alternating potential was used on the grid for the power curve to reduce the effects of distortion, while the maximum permissible crest voltage of 50 volts was used in determining the distortion curve.)

characteristic curves. In an actual tube the minimum plate current is very small, as indicated in Fig. 9-40, but cannot be quite zero.

Effect on Distortion of Changes in Load Resistance. Since it may not always be possible to provide the optimum load resistance for an amplifier, it is desirable to observe the effect on distortion of using a load resistance either higher or lower than the optimum

value. To this end a curve of distortion vs. load resistance in a triode is plotted in Fig. 9-44, with *constant impressed alternating voltage* the crest value of which is just equal to the bias. The curve shows that an increase in load resistance reduces the distortion although, as shown by the accompanying power curve, high values of load resistance cause a considerable decrease in available power output. From this we may conclude that, where it is impossible to supply the optimum load resistance for high output, the load resistance for a triode should preferably be higher, not lower, than the optimum value.

The reason for lower distortion at the higher values of load resistance is that the dynamic characteristic (Fig. 9-35) becomes straighter as the load resistance is increased. That this must be so may be seen from Fig. 9-40 where operation with the 3500-ohm load line is seen to lie much less in the curved region of the static curves than for the 1750-ohm curve for the same grid and plate voltages.

With pentodes and beam tubes the performance is somewhat different, the distortion increasing with a change in load resistance in either direction from the optimum value. The reason for this may be more clearly seen after the discussion of Fig. 9-49 on page 336.

Graphical Determination of Harmonic Content and Power Output with Resistance Loads. Triodes. The exact performance of an amplifier tube, taking into account the curvature of its characteristic curves, may be determined by means of the power series (see Chap. 12), but when the load is a pure resistance the curves of Figs. 9-41 and 9-44 may be computed graphically with a high degree of accuracy and with considerably less work.

The graphical method of determining the amplifier performance requires the drawing of the load line on the static characteristic curves, as in Fig. 9-40. Such a construction is shown in Fig. 9-45 for a load resistance of 3900 ohms and direct voltages of $E_b = 250$ and $E_c = -50$. Let it now be assumed that a potential of $e_g = 50 \sin \omega t$ is applied to the input of the tube. The grid voltage will vary sinusoidally about its average value of -50 volts between the limits of 0 and -100 volts, causing a change in both plate current and plate voltage. Since the plate current and plate potential must follow the load line, the plate current will vary between 64.5 and 3 ma, and the plate potential between 115 and 350 volts. A time

curve of plate current may now be plotted (Fig. 9-46) and analyzed for the presence of harmonics. Table 9-1 shows the result of such analysis up to and including the third harmonic (see Appendix B for the method of solution).

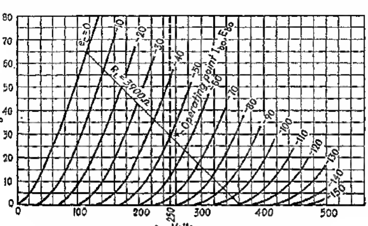


FIG. 9-45. Illustrating the load-line method of amplifier analysis.

FIG. 9-46. Time curve of plate current as obtained from the load line in Fig. 9-45.

Much information may be derived from Fig. 9-45 without the necessity of wave analysis. To facilitate this procedure it will be helpful to redraw the plate current curve of Fig. 9-46 with exaggerated distortion as shown in Fig. 9-47, and write its equation as a Fourier series [see Eq. (B-3) of Appendix B],

$$i_b = a_0 + a_1 \cos \omega t + a_2 \cos 2\omega t + a_3 \cos 3\omega t + \cdots$$
$$+ b_1 \sin \omega t + b_2 \sin 2\omega t + b_3 \sin 3\omega t + \cdots \quad (9\text{-}86)$$

Figure 9-47 shows that the plate-current wave is symmetrical about the t_1 and t_2 axes for $\omega t = 180$ and 360 deg, respectively. This must be so since both increasing and decreasing plate-current values were obtained from the same load line on Fig. 9-45. This symmetry can be expressed by Eq. (9-86) only if each term of that equation is symmetrical about these axes which means that the

TABLE 9-1. RESULTS OF FOURIER ANALYSIS OF THE CURVE OF FIG. 9-46

$R_L = 3900$ ohms, $E_c = -50$ volts, $E_b = 250$ volts

Direct component (with no alternating current) I_{b0}	Direct component (with alternating current) I_b	Fundamental component I_1	Second-harmonic component I_2	Third-harmonic component I_3
29.0	31.8	22.0	1.65	0.16

Currents are all in milliamperes, rms

sine terms must be zero, *i.e.*, b_1, b_2, b_3, etc., are all equal to zero. We may, therefore, rewrite Eq. (9-86) as

$$i_b = a_0 + a_1 \cos \omega t + a_2 \cos 2\omega t + a_3 \cos 3\omega t + \cdots \quad (9\text{-}87)$$

The numerical values of a_0, a_1, etc., may now be determined by first choosing any value of ωt, determining the corresponding value

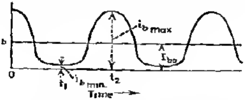


FIG. 9-47. Curve of plate current with exaggerated distortion.

of i_b from Fig. 9-45 and inserting these values of ωt and i_b into Eq. (9-87). If this process is repeated for as many pairs of values of ωt and i_b as there are components to be evaluated, the resulting equations may be solved simultaneously for each component.

Since this method will give satisfactory results only so long as no harmonic of appreciable magnitude is omitted from Eq. (9-87),

it is necessary that the highest order harmonic, which is of sufficient magnitude to be important, be known before a solution is attempted. In triodes only the second harmonic is normally of sufficient magnitude to be considered and we therefore need use only the first three terms of Eq. (9-87); in pentodes and beam tubes the third harmonic is likely to be the largest component and the first four terms of Eq. (9-87) must be used. More terms may be included, if desired, with any type of tube.

Following the general procedure outlined in the preceding paragraphs, equations may be developed by means of which the magnitudes of all the components may be computed directly from curves such as those of Fig. 9-45. Since there are three components to be determined for a triode, *i.e.*, d-c, fundamental, and second harmonic, we must select three different values of ωt and read the corresponding values of i_b from Fig. 9-45. But before we can read i_b we must know the grid voltage, which may be written

$$e_c = E_c + E_{gm} \cos \omega t \tag{9-88}$$

for an impressed sine wave. [Equation (9-88) is written with a cosine rather than a sine term to conform to the zero axis chosen for Fig. 9-47.] With the aid of this equation and the curves of Fig. 9-45 we may now determine i_b for the three values of ωt selected.

Evidently maximum accuracy requires that the three values of i_b be widely different in magnitude; therefore, let us use ωt equal to 0, 90, and 180 deg. At $\omega t = 0$ deg, $\cos \omega t = 1$, $\cos 2\omega t = 1$ and i_b will be at its maximum, so we may write Eq. (9-87) as

$$i_{b\,max} = a_0 + a_1 + a_2 \qquad (\omega t = 0) \tag{9-89}$$

At $\omega t = 180$ deg, $\cos \omega t = -1$, $\cos 2\omega t = 1$, and i_b is at a minimum, so we may write Eq. (9-87) as

$$i_{b\,min} = a_0 - a_1 + a_2 \qquad (\omega t = 180 \text{ deg}) \tag{9-90}$$

At $\omega t = 90$ deg, $\cos \omega t = 0$, $\cos 2\omega t = -1$, and $i_b = I_{b0}$. We may therefore write Eq. (9-87) as

$$I_{b0} = a_0 - a_2 \qquad (\omega t = 90 \text{ deg}) \tag{9-91}$$

monic is plotted to a different scale than the fundamental to show more clearly the relative phase positions

Determination of Harmonic Content and Power Output with Resistance Load. Pentodes and Beam Tubes. The method of the preceding section may be applied to pentodes as shown in Fig 9-49 where the load line represents a resistance of 7000 ohms Since the output of a pentode normally contains more third-harmonic component than second, it is necessary to extend the method to include additional components. This may be done

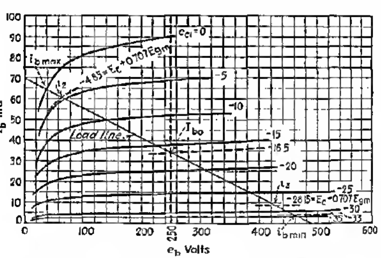


FIG 9-49 Load line for a pentode tube. Resistance load.

by determining the plate current for as many more different grid potentials as there are additional harmonics to be evaluated. To include the third and fourth harmonics let us assume additional grid potentials equal to the grid bias plus and minus 0.707 times the crest value of the alternating grid voltage ($E_c \pm 0.707E_{gm}$), corresponding to values of ωt equal to 45 and 135 deg Letting i_2 and i_3 be the two currents corresponding to these grid voltages (Fig. 9-49), equations may be developed by following the methods of the preceding section, giving

$$I_b = \frac{i_{b\,max} + i_{b\,min} + 2(I_{bo} + i_2 + i_3)}{8} \qquad (9\text{-}103a)^*$$

$$I_{p1} = \frac{i_{b\,max} - i_{b\,min} + \sqrt{2}(i_2 - i_3)}{4\sqrt{2}} \qquad (9\text{-}103b)^*$$

$$I_{p2} = \frac{i_{b\,max} + i_{b\,min} - 2I_{b0}}{4\sqrt{2}} \qquad (9\text{-}103c)^*$$

$$I_{p3} = \frac{i_{b\,max} - i_{b\,min} - \sqrt{2}(i_2 - i_3)}{4\sqrt{2}} \qquad (9\text{-}103d)^*$$

$$I_{p4} = \frac{i_{b\,max} + i_{b\,min} + 2(I_{b0} - i_2 - i_3)}{8\sqrt{2}} \qquad (9\text{-}103e)^*$$

$$\text{Distortion} = \frac{\sqrt{I_{p2}^2 + I_{p3}^2 + I_{p4}^2}}{I_{p1}} \times 100\% \qquad (9\text{-}103f)^*$$

$$P_1 = I_{p1}^2 R_L \qquad (9\text{-}103g)^*$$

Example. The numerical values for substitution in Eqs. (9-103) may be found from Fig. 9-49 for $R_L = 7000$ ohms.

$$I_{b0} = 34,\ i_{b\,max} = 64.5,\ i_{b\,min} = 4,\ i_2 = 61,\ i_3 = 8.5,$$

all in milliamperes.

Substitution of these values into Eqs. (9-103) gives

$$I_b = 34.46 \text{ ma}$$
$$I_{p1} = 23.8 \text{ ma}$$
$$I_{p2} = 0.07 \text{ ma}$$
$$I_{p3} = -2.4 \text{ ma}$$
$$I_{p4} = -0.2 \text{ ma}$$
$$\text{Distortion} = 10.4\%$$
$$\text{Power output} = 3.95 \text{ watts}$$

It should be noted that a crest value of excitation voltage of only 16.5 volts is needed for this pentode tube as compared to the 50 volts required for the triode. This is characteristic of pentodes because of their high μ. It should also be noted that the distortion is somewhat higher than with triodes and is primarily third harmonic in character, which is also characteristic of these tubes when using the load resistance which produces maximum output at low distortion.

The computations of the preceding example may be repeated for other load resistances and the results plotted against R_L as in Fig. 9-50. These curves should be compared with those of Fig.

9-44 for a triode, both sets being computed under the same conditions, i.e., constant impressed signal voltage and bias. It may be seen that for the pentode there is a load resistance that gives minimum distortion, whereas distortion in the triode decreases continuously as the load resistance is increased. The reason for this difference may be seen from Fig. 9-49 which shows that the load line for a pentode extends into the curved region of the characteristic curves at *both* ends instead of at only one end as for the triode (Fig. 9-45). A decrease in load resistance increases the penetration into the distortion region at the lower end, while an

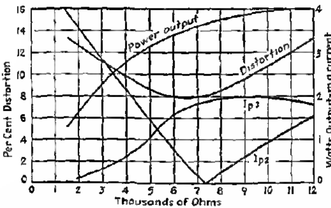


FIG. 9-50 Curves showing results of computations for pentode tube by the method of Fig. 9-49

increase in resistance increases the penetration at the upper end. Thus there is a load resistance for which there is approximately equal penetration at both ends where the second harmonic is zero, as shown in Fig. 9-50, and only the third harmonic remains. Minimum distortion occurs approximately at this point.

It is of interest to note that the second harmonic experiences a 180-deg phase shift as its magnitude passes through zero at approximately 7000 ohms. This is because the plate-current wave will be flattened in the region of *minimum* current for resistances less than about 7000 ohms but in the region of *maximum* current for higher resistances.

Equations (9-103) are equally applicable to a triode and should be used when there is a probability of appreciable harmonic output of higher order than the second, as in the case of a badly overloaded amplifier. They *must* be used for pentodes and beam tubes, since, as illustrated in the pentode example, the third harmonic may be larger than the second.

Application of this analysis to beam tubes follows the pentode example just presented and produces very similar results.

Summary. The performance of a power amplifier with resistance load may be predicted from the construction of Figs. 9-45 or 9-49 and from Eqs. (9-101) and (9-102) or Eqs. (9-103*f*) and (9-103*g*) for any given load resistance and direct plate and grid voltages.[1] The correct grid bias for any given plate voltage may be obtained by the graphical method of Fig. 9-40. Under normal conditions, however, the correct grid and plate voltages, the optimum value of load resistance, the power output, and the distortion are obtainable from tube data sheets provided by the tube manufacturers, Eqs. (9-101) to (9-103) providing information for any other voltages and load resistances as desired.

Effect of Reactive Loads. The analysis of the preceding sections is applicable only to *resistive* loads. If the load contains some reactance, the load curve is no longer a straight line but an ellipse (Fig. 9-51). An analysis of this type of load curve is extremely difficult, and the power-series analysis of Chap. 12 should be used.

An example of reactive load is found in loud-speakers which may be represented by an equivalent circuit containing both resistance and *reactance*. Although computations of sufficient accuracy for many purposes may be carried out by considering such a load impedance to be resistive only, the reactive component

[1] Some error is involved in this method because the assumption of a pure resistance load means that the increase in the direct component of plate current due to rectification, $(I_b - I_{bo})$, should encounter the same resistance in the plate circuit as do the alternating components. With transformer coupling this is not so, and the error involved may be appreciable unless the rectified current is small. For details, see C. E. Kilgour, Graphical Analysis of Output Tube Performance, *Proc. IRE*, 19, p. 42, January, 1931.

The power series method of analysis (Chap. 12) does not include this error.

of the load may produce resonance and other related phenomena and must therefore be carefully considered in any thorough evaluation of amplifier performance.[1]

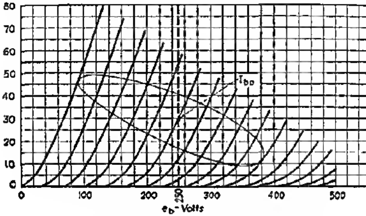


FIG. 9-51 Load line of a triode with reactive load

Plate Efficiency. Triode Tubes. The plate efficiency of an amplifier is the ratio of the a-c power output to the d-c power input to the plate circuit, or

$$\text{Plate efficiency} = \frac{P_o}{E_b I_b} = \frac{E_p I_p}{E_b I_b} \qquad (9\text{-}101)$$

where P_o is the power output of the tube. In the triode example of page 333 this equation gives

$$\text{Plate efficiency} = \frac{1.84}{250 \times 31.4 \times 10^{-3}} \times 100 = 23.4\%$$

The maximum theoretical efficiency of a class A amplifier is 50 per cent. This may be seen by referring to Fig. 9-52, which shows a set of static characteristic curves and a load line *ab*. The slope of the load line is determined by the load resistance of the amplifier, as previously pointed out in connection with Fig 9-40

[1] An analysis of amplifiers with reactive load is given by Manfred von Ardenne, On the Theory of Power Amplification, *Proc. IRE*, 16, p 193, February, 1928.

Evidently the theoretically maximum possible swing of the plate voltage is from point a to point b, or from 0 to $2E_b$,[1] where E_b is made equal to half the voltage $0b$. The plate voltage cannot swing farther toward the left, as it is impossible for it to become negative; it cannot swing farther to the right, as this would require a negative plate current, also an impossibility. Moreover, if the plate voltage is varied from 0 to b, the plate current will also vary over its maximum possible range, from a to 0 (*i.e.*, from $2I_b$ to 0).

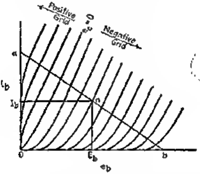


FIG. 9-52. Illustrating the maximum theoretical values of alternating plate voltage and current in a triode.

Therefore, *under these maximum theoretical conditions*, we may write $E_{pm} = E_b$, and $I_{pm} = I_b$, or $E_p = 0.707E_b$, and $I_p = 0.707I_b$. Substituting in Eq. (9-104) gives

$$\text{Max theoretical plate eff.} = \frac{(0.707E_b)(0.707I_b)}{E_bI_b} \times 100 = 50\%$$

Actually the efficiency of any class A amplifier will be much less than its theoretical maximum of 50 per cent. An inspection of Fig. 9-52 will show that the grid must swing highly positive to make the plate voltage even approach zero. Since a highly positive grid swing will produce considerable distortion, it is not practical to make E_{pm} equal to E_b as was assumed. Also, if the plate current is to swing to zero, the lower values of current will lie in the

[1] Such a large variation is impossible physically, since it calls for *maximum* plate current with *zero* plate voltage, but represents the maximum *theoretical* value.

curved region of the static curves; therefore, the output wave will be flattened during both halves of the grid-voltage swing. Neither the plate voltage nor the plate current can, therefore, ever swing to zero, and the efficiency is always less than 50 per cent, 25 per cent being about the maximum, as in the preceding example.

The *average* efficiency of a class A amplifier, averaged over a period of minutes, is always very much less than that computed from Eq. (9-101). This is because in normal operation the excitation voltage varies from a maximum to zero, as for example in the amplification of speech or music where maximum output is obtained only at the highest sound levels. Therefore the a-c output varies from a maximum to zero, with the average value very much less than the maximum, while the input remains virtually constant, since E_b does not vary and I_b varies only slightly as the excitation is changed. The efficiency is proportional to the output under such conditions, and the average efficiency will be only a few per cent.

The power loss on the plate of a class A amplifier is a maximum when the excitation voltage and output are zero. Since the power input, E_bI_b, is nearly constant in a class A amplifier regardless of the magnitude of the excitation voltage and the plate loss is equal to the difference between the power input and the power output, the loss must *increase* as the excitation (and therefore the power output) *decreases*, being equal to the power input when the excitation voltage is zero. Thus the tube must be capable of dissipating a maximum amount of power equal to E_bI_b.

Plate Efficiency. Pentode and Beam Tubes. The foregoing analysis of the efficiency of class A amplifiers applies equally well whether triode, pentode, or beam tubes are used, but the efficiency of pentode and beam tubes is somewhat higher than that of triodes. This is because the minimum plate voltage in a pentode (or beam tube) is much lower than in a triode which makes E_p larger, thus increasing the ratio of Eq. (9-101). This is illustrated by Figs. 9-45 and 9-49 where the minimum plate voltage of the triode is 117 while that of the pentode is 40. Since E_b for both tubes is 250 volts, E_p for the triode is $0.707\ (250 - 117) = 94$ volts and for the pentode $0.707\ (250 - 40) = 148.5$ volts. A lower minimum plate voltage is possible in pentodes because the electrons are drawn away from the cathode by the *constant-potential* screen grid, and the plate potential need be only sufficient to collect the elec-

trons that pass through the screen, while the plate of the triode must remain sufficiently positive to attract the necessary number of electrons away from the cathode.

The plate efficiency of the pentode amplifier of the example on page 335 is found from Eq. (9-101) to be

$$\text{Plate efficiency} = \frac{3.95}{250 \times 34.45 \times 10^{-3}} \times 100 = 46\%$$

The over-all efficiency is somewhat less due to the screen-grid loss. Adding $E_{c2}I_{c2}$ to E_bI_b to obtain the total input, where $E_{c2} = 250$ volts and $I_{c2} = 6.5$ ma, gives

$$\text{Plate efficiency} = \frac{3.95}{8.61 + 1.02} \times 100 = 38.6\%$$

Push-pull Amplifiers. Most power amplifiers are now constructed with two tubes in push-pull (Fig. 9-53) and are operated class A_1, class A_2, class AB_2, or class B_2.[1] It is obvious that the emfs applied to the grids of the tubes by the two halves of the

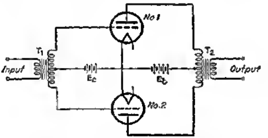


FIG. 9-53. Circuit of a push-pull amplifier.

input transformer are of opposite polarity. One tube is, therefore, operating on the upper portion of its characteristic curve while the other is operating on the lower portion of its curve, or, in the case of class AB and class B amplification, its plate current may even be zero. In the output transformer T_2 the alternating components are again combined with proper polarity, and an amplified signal is delivered to the output terminals.

The principal advantages of push-pull amplification, when the

[1] See footnote in Appendix A (p. 602) for the meaning of the subscripts 1 and 2.

two tubes have identical characteristics, are: (1) The even harmonics are balanced out in the output circuit and thus a greater output may be obtained without exceeding a given permissible distortion; (2) hum voltages introduced by the plate-supply source will balance out; (3) d-c saturation of the output transformer is avoided, since the two tubes draw direct plate current through the two halves of the primary winding in opposite directions; and (4) no signal-frequency component of current flows through the plate-supply source to produce feed-back problems. For these four reasons most power amplifiers and even many voltage amplifiers are constructed in push-pull.

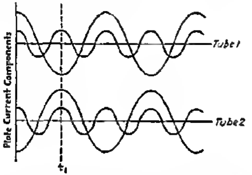


FIG. 9-54. Plate current in each tube of Fig. 9-53.

That the even harmonics balance out may be seen from Fig. 9-54. Here the curves of Fig. 9-48 are reproduced for tube 1 and similar curves are drawn for tube 2 except that they are displaced 180 deg since the grid voltage on tube 2 is 180 deg out of phase with that applied to tube 1. In the output transformer T_2 of Fig. 9-53 the two tube currents oppose each other in producing ampere turns in the primary winding. This means that, if the two tubes are exactly alike, the fundamental components of the two currents in Fig. 9-54 will produce additive magnetization in the transformer core, but the two second-harmonic components will cancel each other. If this analysis were carried to higher order harmonics, it would be seen that all even harmonics would produce zero net magnetization of the transformer core while all odd harmonics would be additive.

It is of interest to note that in the common return for the plate current of the two tubes, through the d-c source E_b, all even harmonics of the currents in the two tubes are additive but all odd harmonics cancel out.

Class A Push-pull Amplifier Performance.[1] It might at first be supposed that each tube of Fig. 9-53 is working into a load resistance equal to the resistance of the output circuit times the square of the turn ratio between one-half of the primary and the secondary, but this neglects the reaction of the other tube. The situation may be made clear by reference to Fig. 9-55 which shows two alternators supplying a load R_L through a mid-tapped trans-

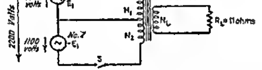


FIG. 9-55. Illustrating the action of a push-pull amplifier by means of two alternators supplying a single load through a mid-tapped transformer.

former. The two alternators have equal voltages and are exactly in phase considering the series circuit; they are therefore equivalent to the two tubes of the push-pull amplifier. Suppose that the terminal voltage of each alternator is 1100 volts and that the transformer is 2200/1100 to 110 volts and that R_L is 11 ohms. If switch S is open so that alternator 2 is disconnected, the secondary load current will be 10 amp and the current through alternator 1 will be 1 amp. The effective primary resistance presented to alternator 1 will be $1100/1 = 1100$ ohms, which is $(N_1/N_L)^2 \times R_L$, where N_1 is the number of turns in the upper half of the primary and N_L is the number of turns in the secondary. The power delivered will be $1100 \times 1 = 1100$ watts, which is of course equal to the power absorbed in R_L.

Now suppose the switch S to be closed. The secondary emf will obviously be the same as before; and therefore the secondary

[1] See also B. J. Thompson, Graphical Determination of Performance of Push-pull Audio Amplifiers, *Proc. IRE*, **21**, p. 591, April, 1933; and Albert Preisman, Balanced Amplifiers, *Communication and Broadcast Eng.*, **3**, p. 12, February, 1936.

current will remain at 10 amp, and the total power output at 1100 watts. No current will flow through the middle wire on the primary side, however, so that the primary current must flow through the two alternators in series, having a magnitude of 0.5 amp. The total power delivered will remain the same as before switch S was closed, *viz.*, 1100 watts, and each alternator will deliver half, or 550 watts. In other words, connecting the second alternator into the circuit reduced the power demanded of alternator 1 but did not affect the load at all. The effective primary resistance between outside leads is now $2200/0.5 = 4400$ ohms, or that across the terminals of one alternator is 2200 ohms, twice the resistance presented when only one alternator was in operation. This rise in the effective resistance across the terminals of alternator 1 results from the reaction produced by alternator 2 through the mutual induction between the two halves of the primary winding. If the two alternators are to produce twice the power output of the one, it is evidently necessary that the load resistance used with the two alternators be but half that used with the one.

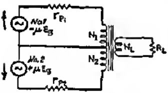


FIG. 9-56. Equivalent a-c circuit of a push-pull amplifier. Transformer ratio $N_1/N_L = N_2/N_L = 1$

The equivalent circuit of a push-pull amplifier is shown in Fig. 9-56 where the ratio of total primary to secondary turns is assumed to be 2; *i.e.*, if one tube is removed, the other will work into a load of R_L through a 1:1 transformer. Applying the results of the preceding paragraph, it is evident that, for double the power output of one tube to be obtained, the load resistance R_L must be half that used with only one tube coupled through its half of the primary winding. Actually it may be made even lower with triodes, such that each tube will work into an impedance of approximately r_p, instead of $2r_p$ as in the *single-tube amplifier*. *That this is possible without* excessive distortion may be seen from the graphical analysis presented in the following section.

Composite Characteristic Curves. The operating characteristics of push-pull amplifiers may be determined graphically[1] from the characteristic curves of the tubes in much the same manner as

[1] For a more complete discussion of this method, see Thompson, *loc. cit.*

for the single-tube amplifier. The characteristic curves of both tubes must be obtained and plotted as in Fig. 9-57 with those of one tube inverted. The curves must be so plotted that the normal

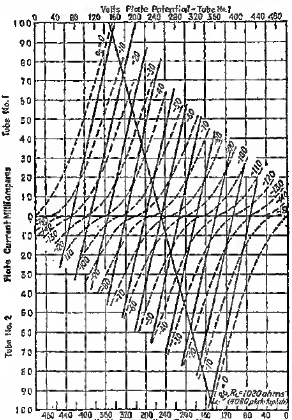


FIG. 9-57. Static characteristic curves of two triode tubes shown by the broken lines, with composite curves for push-pull operation shown by solid lines.

plate voltage—in this case 250 volts—coincides for the two characteristics. It is then possible to plot composite characteristic curves for the two tubes acting in series across the total primary winding. This is done for the normal grid-bias curves—in the

problem at hand -50 volts—by merely adding the two tube currents algebraically, which will give the net current producing magnetization in the core of the transformer. For other grid voltages it is necessary to combine curves that are equally above and below the normal grid bias, since the two tubes operate with opposite polarity of alternating input voltage on their grids. For example, the curve for -40 volts (10 volts above the normal bias) on tube 1 is combined with that for -60 volts (10 volts below the normal bias) for tube 2. If the characteristics of the two tubes are similar, the composite curves will be virtually straight lines and there will therefore be but little distortion.[1]

These composite curves represent the characteristics of an equivalent tube which, if used in place of one of the two push-pull tubes with the other removed, would produce the same distortion and output as are produced by the actual push-pull amplifier. The plate resistance $r_{p_{eq}}$ of this composite tube may be determined by measuring the slope of the curves, being about 1020 ohms, or a little more than half that of the actual tubes.

Theoretical Determination of the Correct Load Resistance for Push-pull Amplifiers. For push-pull amplifiers the theoretically correct load resistance is $R_L = r_{p_{eq}}$, not $R_L = 2r_{p_{eq}}$ as for single-tube amplifiers. To demonstrate this we may follow the same method of approach used with Fig. 9-42, page 324, drawing the theoretical straight-line characteristic curves of Fig. 9-58 for the equivalent tube to be used in place of tube 1, Fig. 9-53, with tube 2 removed. Here operation is along a load line EDE' which has for its limits two parallel lines, $0E$ and $0'E'$, instead of two intersecting lines as in the single-tube construction of Fig. 9-42.

We may now apply Eqs. (9-77) to (9-79), page 324, to the equivalent tube of the push-pull amplifier. In this equivalent tube $r_{p_{eq}} = 0A/AE = 0A/I_{pm}$ (not $0A/2I_{pm}$ as in the single-tube amplifier) which means that Eq. (9-79) becomes, for the push-pull amplifier,

$$E_{pm} = E_b - I_{pm}r_{p_{eq}} \tag{9-105}$$

[1] Note that the solid curves of Fig. 9-57, obtained in this manner, represent the equivalent of a constant grid voltage but with a varying potential between the anodes of the two tubes such that the average anode potential of the two is at all times equal to E_b.

Substituting this equation into Eq. (9-78) and solving for I_{pm} gives

$$I_{pm} = \frac{E_b}{R_L + r_{p_{eq}}} \qquad (9\text{-}106)$$

When this equation is substituted into Eq. (9-77), the result is

$$P = \frac{E_b^2 R_L}{2(R_L + r_{p_{eq}})^2} \qquad (9\text{-}107)$$

This equation may now be differentiated with respect to R_L and the result set equal to zero to give

$$R_L = r_{p_{eq}} \qquad (9\text{-}108)^*$$

This is in contrast to the single-tube amplifier where maximum undistorted power output was obtained when $R_L = 2r_p$.

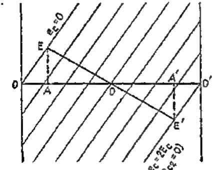


FIG. 9-58. Theoretical straight-line characteristic curves for a push-pull amplifier.

Graphical Solution of Push-pull Amplifier Performance. A graphical solution of the performance of the push-pull amplifier may be obtained by following the same procedure as was followed for the single-tube amplifier. Load lines may be drawn exactly as was done for the single-tube amplifier, the one shown in Fig. 9-57 being for the optimum resistance of $R_L = r_{p_{eq}}$, or 1020 ohms, across a secondary having the same number of turns as one-half the primary, to give a load resistance of 1020 ohms across the terminals of the equivalent tube. The plate-to-plate load resistance of the actual push-pull amplifier will be four times this

amount, or 4080 ohms, so that the load resistance per tube is 2040 ohms. Half this resistance, as already pointed out, is due to the reaction of the 1020-ohm load, and half to the effect of the other tube acting through the mutual inductance of the transformer windings.

The power output may be determined by applying Eq. (9·103g) to the equivalent tube. For $E_{gm} = E_c$; $i_{b\,max} = 0.095$, $i_{b\,min} = -0.095$, $i_2 = 0.001$, $i_3 = -0.061$, and $I_{b0} = 0$. Using these values in Eq. (9·103g) gives $P = 4.4$ watts, nearly three times, rather than twice, the maximum power for the single tube as computed on page 333. This greater power output is due to the use of a resistance approximately equal to r_p across each tube instead of $2r_p$.[1]

The foregoing discussion of push-pull, class A amplifiers has been confined to triode tubes, but pentodes and beam tubes are also widely used in push-pull. The performance of an amplifier using pentode or beam tubes may be determined in the same manner as that of a triode amplifier, although, since the even harmonics are nearly nonexistent in properly designed, single-tube pentode amplifiers, the use of a push-pull circuit does not result in so great an improvement in performance as with triodes. Nevertheless the reduction in hum, the elimination of transformer saturation, and other factors provide sufficient improvement in performance to make the push-pull circuit attractive for pentode tubes.

Class B Push-pull Amplifier.[2] Two tubes operating push-pull, class B, may be made to operate virtually as a class A amplifier, producing an output voltage reasonably proportional to the input voltage at all frequencies within the band for which the amplifier

[1] This output is nearly twice the output of 2.3 watts computed for one of these tubes working into a load resistance of r_p with $E_{gm} = E_c$ and *assuming no distortion*. If no distortion were present here and if r_{peq} were exactly half the r_p of each individual tube, the output would be exactly twice 2.3.

[2] See also Loy E. Barton, High Audio Power from Relatively Small Tubes, *Proc. IRE*, **19**, p. 1131, July, 1931; Application of the Class B Audio Amplifier to A-c Operated Receivers, *Proc. IRE*, **20**, p. 1085, July, 1932, Recent Developments of the Class B Audio- and Radio-frequency Amplifiers, *Proc. IRE*, **24**, p. 985, July, 1936; J. R. Nelson, Class B Amplifiers Considered from the Conventional Class A Standpoint, *Proc. IRE*, **21**, p. 858, June, 1933, True McLean, An Analysis of Distortion in Class B Audio Amplifiers, *Proc. IRE*, **24**, p. 487, March, 1936; Glenn Koehler, Class B and AB Audio Amplifiers, *Electronics*, **9**, p. 14, February, 1936.

is designed. The circuit used is exactly the same as that for the class A amplifier (Fig. 9-53), but the tubes are biased approximately to cutoff.

The operation of a class B push-pull amplifier may be best explained by reference to Fig. 9-59. Here the dynamic characteristics of the two tubes are plotted back to back in such a manner that the grid-bias voltages for the two tubes coincide at the point p. Consequently the excitation voltage may be shown as a sine wave superimposed upon a vertical line passing through the point p. In this way the current flowing in the plate circuit of either

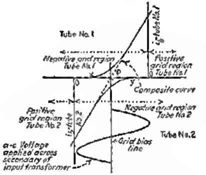


FIG. 9-59. Individual and composite dynamic characteristic curves of two tubes operating in push-pull, class B.

tube at any given instant of time may be determined by projecting a vertical line upward from the value of the instantaneous voltage, as given by the sine wave, until it intersects the characteristic curves of the tubes. Evidently only one tube will be carrying current at any given time except for the short interval xy, where the net current flowing, so far as magnetization of the transformer core is concerned, will be the difference and, therefore, be represented by the dotted line. It will be observed that this dotted line together with the characteristic curves lying outside the xy region constitutes a nearly straight line which is more than twice as long as the straight-line portion of the characteristic curve of either tube alone. It will also be observed that the correct grid bias is that at which the plate current in each tube would become zero if the substantially straight portion of its curve were projected to the zero line.

If the theoretically straight line of Fig. 9-59 could be exactly obtained in practice, class B push-pull amplifiers would give distortionless amplification, but such a condition is difficult of attainment. In the first place, the characteristic curves of each individual tube are not straight, even at the higher plate currents, so that the composite curve tends to be slightly S-shaped. Furthermore, any slight variation in the grid bias or plate voltage is equivalent to sliding the two characteristic curves relative to one another. If this change is in the nature of an increase in the negative grid bias or a decrease in the plate voltage, the curves of Fig. 9-59 will

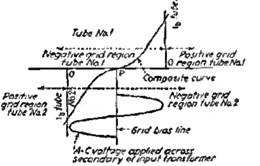


FIG. 9-60 Showing the effect on the performance of a class B push-pull amplifier of a bias that is too negative.

appear as in Fig. 9-60 where the resulting dynamic curve is far from straight, and serious distortion will result. It is, therefore, evident that a very constant source of direct potentials must be provided or high-mu tubes, designed for operation with zero bias, must be used.

Effect of Driving Grid Positive. The maximum capacity of the class B push-pull amplifier is realized only when the grid is swung over into the positive region.[1] This, of course, means grid-current flow and, as we have already seen, a tendency toward distortion This distortion is caused, however, not by the mere fact that grid current flows but because the grid current in the average tube

[1] With zero-bias tubes there could, of course, be no output without operation in the positive-grid region.

does not vary in direct proportion to the applied potential; *i.e.*, the input impedance varies throughout the cycle.[1] This change in impedance causes a much higher drop through the input or driver circuit at certain times in the cycle than at others, thus producing distortion. The distortion may, therefore, be made very small by either (1) using a low-impedance driver, so that the drop will be almost negligible, or (2) maintaining the input impedance to the class B stage reasonably constant throughout the cycle so that the drop will be directly proportional to the applied signal.

A low-impedance driver is generally secured by using a stepdown transformer between the preceding, or driving, tube and the grids of the class B stage. The plate resistance of the driver tube, as seen from the secondary side of the transformer, is therefore low. Another way of stating this is that the load impedance on the driving tube is made quite high through the action of the transformer which steps up the input resistance of the class B stage as seen from the driver, or input, side of the transformer. By combining Eqs. (3-20) and (3-21) (pages 63 and 64), we may write

$$E_p = -\mu E_e \frac{R_L}{r_p + R_L} \tag{9-109}$$

which shows that the driver plate voltage (and, therefore, the grid voltage of the succeeding class B stage) is relatively independent of variations in R_L so long as R_L is large compared to r_p. Of course, the power capacity of the driver stage is reduced when R_L is larger than r_p, but the power demand, although greater than that of a class A_1 stage, is not large. The distortion in the driver stage is also reduced by an increase in R_L (if a triode is used) as shown by the curve of Fig. 9-44.

Tubes designed for zero grid bias tend to have a reasonably constant input impedance. The dynamic curves of a pair of such tubes are shown in Fig. 9-61 where zero grid potential is the common point p. While these tubes draw grid current throughout the entire positive half cycle of the signal voltage, the grid-current–grid-voltage curve is linear up to a fairly high positive grid potential as shown in Fig. 9-61. This means that the input impedance to a pair of these tubes in push-pull is practically constant throughout the region uv, and the distortion resulting from grid-current flow

[1] See also the discussion on p. 319 in connection with Fig. 9-37.

is very small so long as the crest value of the excitation voltage does not exceed *up* (or *pv*).

As previously stated, the power required to drive this type of amplifier is somewhat greater than that required by class A_1 amplifiers where the input impedance is extremely high (grid current practically zero). Consequently the driver stage (the amplifier that supplies the excitation voltage) must have sufficient power capacity to handle this higher demand. The driving power necessary for tubes of a given type is usually given in tube handbooks or may be determined from the tube-characteristic curves, and the driver stage must be designed to deliver this power plus any losses in the coupling network.

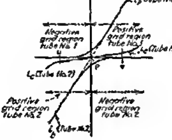


Fig. 9.61 Individual and composite dynamic characteristic curves of two zero-bias tubes in push-pull, class B

Class AB Push-pull Amplifier. The class AB amplifier operates with a bias somewhere between that required for class A and that required for class B amplification. Thus it is less subject than the class B amplifier to increased distortion with changes in direct voltage supply, while possessing much of the increased power capacity of the class B, and is probably more widely used for power-amplification purposes than are either class A or class B amplifiers. As in the case of class B amplifiers the grids of the tubes are ordinarily driven somewhat positive to secure the maximum power output.

Determination of Class B and AB Push-pull Amplifier Performance. The performance of class B and AB, push-pull amplifiers may be determined in exactly the same manner as for class A

amplifiers.[1] Figure 9-62 shows a set of curves for a pair of tubes operating with a 65-volt negative bias and 250 volts on the plate. The composite curves are drawn in the manner already described

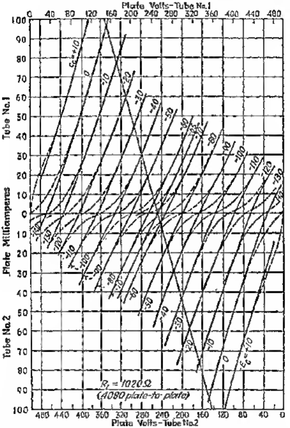


FIG. 9-62. Construction for determining the operating characteristics of two tubes in push-pull, class B.

in connection with Fig. 9-57, and load lines may be drawn for any desired resistance. The one shown is for 4080 ohms, plate to

[1] B. J. Thompson, Graphical Determination of Performance of Push-pull Audio Amplifiers, *Proc. IRE*, **21**, p. 591, April, 1933.

plate. The performance may be determined in the same manner as for the class A amplifier, but the normal range of operation extends over into the positive grid region.[1]

Plate Efficiency of Push-pull Amplifiers. The plate efficiency of class A push-pull amplifiers is slightly higher than that of single-tube amplifiers (see page 338). The reason for this is that a smaller minimum plate current is permissible due to cancellation of the even harmonics. Thus operating conditions are a little nearer the theoretical maximum given on page 339.

Class B push-pull amplifiers have a much higher efficiency, generally of the order of 60 to 65 per cent, class AB amplifiers have an efficiency lying between that of class A and class B amplifiers,

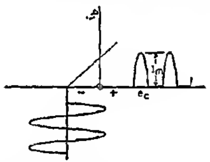


FIG. 9-63. Wave shape of the plate current in a class B amplifier with idealized characteristic.

probably not to exceed about 50 to 55 per cent and in many cases somewhat lower.

As pointed out on page 340, the efficiency when using pentode tubes is a little higher than when using triodes, although the improvement in the case of class B amplifiers is less than for class A amplifiers.

The maximum theoretical efficiency of a class B amplifier may be found by assuming the dynamic characteristic to be a straight line. By the definition of a class B amplifier (Appendix A) the plate current will then consist of half sine-wave pulses (Fig. 9-63).

[1] An excellent method of analyzing class *B* amplifiers is that given by W. G. Wagener, Simplified Methods for Computing Performance of Transmitting Tubes, *Proc. IRE*, **26**, p. 47, January, 1937.

Analyzing this plate-current wave by the method of Appendix C gives, for a sine wave of impressed voltage,

$$i_b = 0.318 I_m + 0.5 I_m \sin \omega t - 0.212 I_m \cos 2\omega t + \cdots \qquad (9\text{-}110)$$

where I_m is the peak value of the plate-current pulse as indicated in Fig. 9-63. From Eq. (9-110) we may write

$$I_b = 0.318 I_m \qquad \text{and} \qquad I_p = \frac{0.5 I_m}{\sqrt{2}}$$

where I_p is the rms value of the fundamental component of the plate current.

The plate voltage in a properly designed class B push-pull amplifier is essentially sinusoidal, and the maximum theoretical alternating plate voltage is given by

$$E_{pm} = E_b \qquad \text{or} \qquad E_p = \frac{E_b}{\sqrt{2}}$$

which corresponds to the assumptions made in determining the maximum theoretical plate efficiency of the triode (page 338).

Using the foregoing relations the maximum theoretical plate efficiency may be written from Eq. (9-104) (page 338) as

$$\text{Max theoretical plate efficiency} = \frac{E_b(0.5 I_m) \times 100}{\sqrt{2}\sqrt{2} E_b (0.318 I_m)} = 78.5\%$$

The crest value of the plate voltage of a practical class B amplifier is a greater percentage of E_b than is that of a class A amplifier, but it can never reach E_b (*i.e.*, $E_{pm} < E_b$). The maximum efficiency of a practical amplifier is therefore less than 78.5 per cent, being about 60 to 65 per cent.

The *average* efficiency of a class B amplifier is much closer to the maximum than is that of a class A amplifier. This is because the input power as well as the output varies with the excitation voltage. If both were directly proportional to the excitation, the efficiency would be constant for all values of signal voltage but, owing to the curvature of the characteristic curves, the input power decreases less rapidly than the output at low signal voltages, being somewhat greater than zero at zero excitation. The average efficiency of class AB amplifiers is appreciably less than that of class B amplifiers since the no-signal plate current is higher.

Phase-inverting Tube to Drive a Push-pull Amplifier. Although

push-pull amplifiers may be driven through an input transformer as in Fig. 9-53, it is usually preferable to drive them by a resistance-coupled amplifier and phase-inverting tube to obtain the better frequency response and lower initial cost of a resistance-coupled amplifier. A commonly used circuit of this type is shown in Fig. 9-64. Tubes 3 and 4 are the push-pull amplifier tubes which are coupled to the output circuit in the usual manner through transformer T. The input signal is applied to the grid of tube 1, and the output of this tube drives the grid of tube 3 through the usual resistance-coupled network. A portion of this output voltage is taken off to drive the grid of tube 2, and the output of this tube in turn drives the grid of the other push-pull amplifier tube 4. Since the plate voltage of a vacuum-tube amplifier with resistance load is out of phase with the grid voltage by 180 deg, it is evident that the grids of tubes 3 and 4 are excited in phase opposition, as they should be for push-pull operation. It is further evident that

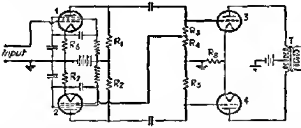


FIG. 9-64. Circuit of a resistance-coupled amplifier and phase inverter, driving a push-pull amplifier.

the output voltage of tube 2 should be equal to that of tube 1; therefore, the reduction in voltage through the resistances R_3 and R_4 should be just sufficient to offset the gain of tube 2.

Figure 9-65 shows another commonly used circuit wherein a resistor in series with the cathode of tube 1 is used to provide the excitation for the second tube of the push-pull amplifier. Since the same alternating current flows through both R_1 and R_2 (being the plate current of tube 1) the only requirement for securing equal voltages on the grids of tubes 2 and 3 is that the two resistances be equal. The voltage across R_2 is also applied to the grid of tube 1, producing negative feedback (see material beginning

on page 361) and resulting in an over-all gain for tube 1 of slightly less than two.

A number of other circuits are used for phase inversion but the two presented here are typical.[1]

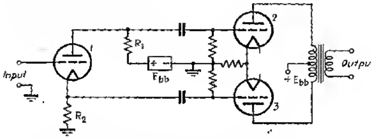


FIG. 9-65. Phase inverter using a cathode resistor.

Relative Advantages of Class B and AB Amplifiers and Class A Amplifiers. The principal advantage of class B and class AB amplifiers over class A amplifiers lies in their increased power capacity and their higher average efficiency. The increased power capacity is due largely to the higher efficiency obtainable which permits more power output without overheating of the tubes. Since distortion, rather than overheating, may impose the limit on output, increased capacity may also be due to the greater permissible plate-current swing or to the larger excitation voltage which may be applied because of the higher bias used.

A principal disadvantage of class B and AB amplifiers over class A amplifiers is that the distortion produced at normal output is higher. Thus class A push-pull amplifiers are sometimes used as power amplifiers where high quality is desired, whereas class AB amplifiers are used in most applications.

Frequency Distortion in Power Amplifiers. All reference to distortion in the preceding sections on power amplifiers has been with respect to amplitude distortion, as this is the type most likely to appear in power amplifiers. Frequency distortion may also be present to some degree for the same general reasons as were presented under voltage amplifiers; *i.e.*, the shunting effect of the interelectrode capacitances of the tubes and the variations with frequency of the impedances of the circuit elements. Since vir-

[1] Other circuits are given by Myron S. Wheeler, An Analysis of Three Self-balancing Phase Inverters, *Proc. IRE*, **34**, p. 67P, February, 1946.

tually all power amplifiers are transformer-coupled to the load, the circuit to be analyzed is that of Fig. 9-19, page 295, except that a resistance $R_L = R_s/n^2$ must be connected across the terminals GK to represent the load R_s of Fig. 9-39 (where n is the turns ratio of

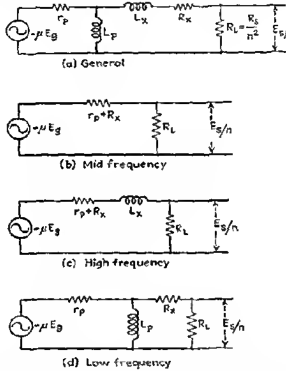


FIG. 9-66. Equivalent circuits of a transformer-coupled power amplifier.

secondary to primary and is usually less than unity).[1] This resistance is normally very much less than the reactance of C_T at any frequency for which the gain of the amplifier is reasonably close to the mid-frequency gain and C_T may be omitted. Thus

[1] While this analysis is applied to a single-tube amplifier, it may also be applied to a push-pull amplifier by replacing the push-pull circuit with that of the equivalent tube referred to on page 316.

the equivalent circuit corresponding to Fig. 9-19b is shown in Fig. 9-66a, with the mid-, high-, and low-frequency circuits shown in b, c, and d, respectively.

The mid-frequency gain may be readily found from Fig. 9-66b by first writing the ratio

$$\frac{\mathbf{E}_s/n}{-\mu\mathbf{E}_g} = \frac{R_L}{r_p + R_x + R_L} \tag{9-111}$$

or

$$\mathbf{A}_m = \frac{\mathbf{E}_s}{\mathbf{E}_g} = \mu n \frac{R_L}{r_p + R_x + R_L} \underline{/180^\circ} \tag{9-112*}$$

Similarly the h-f gain may be found from Fig. 9-66c by writing

$$\frac{\mathbf{E}_s/n}{-\mu\mathbf{E}_g} = \frac{R_L}{(r_p + R_x + jX_x) + R_L} \tag{9-113}$$

where $X_x = \omega L_x$. We may now write from Eq. (9-113),

$$\mathbf{A}_h = \frac{\mathbf{E}_s}{\mathbf{E}_g} = \mu n \frac{R_L}{r_p + R_x + R_L} \frac{1}{\sqrt{1 + \dfrac{X_x^2}{(r_p + R_x + R_L)^2}}} \underline{\Big/ 180^\circ - \tan^{-1} \frac{X_x}{r_p + R_x + R_L}} \tag{9-114*}$$

For the l-f response the circuit of Fig. 9-66d may be simplified by applying Thévenin's theorem to that part of the circuit lying

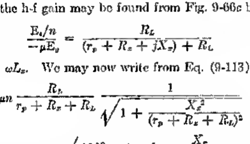


FIG. 9-67. Simplification of the circuit of Fig. 9-66d by means of Thévenin's theorem.

to the left of R_L, to obtain the circuit of Fig. 9-67. We may write the ratio

$$\frac{\mathbf{E}_s/n}{-\mu\mathbf{E}_g \dfrac{jX_p}{r_p + jX_p}} = \frac{R_L}{R_L + \left(R_x + \dfrac{jX_p r_p}{r_p + jX_p}\right)} \tag{9-115}$$

where $X_p = \omega L_p$. Equation (9-115) may be rewritten as

$$\mathbf{A}_l = \frac{\mathbf{E}_2}{\mathbf{E}_g} = -\mu n \frac{X_p R_L}{X_p(r_p + R_e + R_L) - jr_p(R_e + R_L)} \qquad (9\text{-}116)$$

This may be rearranged and changed to polar form to give

$$\mathbf{A}_l = \mu n \frac{R_L}{r_p + R_e + R_L} \frac{1}{\sqrt{1 + \left[\frac{r_p(R_e + R_L)}{(r_p + R_e + R_L)X_p}\right]^2}} \Big/ 180^\circ + \tan^{-1} \frac{r_p(R_e + R_L)}{(r_p + R_e + R_L)X_p} \qquad (9\text{-}117)^*$$

Equation (9-117) gives results that are slightly in error if the resistance of the transformer windings is appreciable compared to r_p and R_L. From the circuits of Figs. 9-18 and 9-19 it may be seen that R_e is made up of the resistance of the primary, R_1, and the effective resistance of the secondary, R_2/n^2, and that R_1 actually should be in series with r_p in Fig. 9-66*d* while the series resistance to the right of L_p should be R_2/n^2, not R_e. The effect of this correction on Eq. (9-117) is to increase r_p to $r_p + R_1$ and to replace R_e with R_2/n^2. Making these changes gives, for the l-f gain

$$\mathbf{A}_l = \mu n \frac{R_L}{r_p + R_e + R_L} \frac{1}{\sqrt{1 + \left[\frac{(r_p + R_1)(R_L + R_2/n^2)}{(r_p + R_e + R_L)X_p}\right]^2}} \Big/ 180^\circ + \tan^{-1} \frac{(r_p + R_1)(R_L + R_2/n^2)}{(r_p + R_e + R_L)X_p} \qquad (9\text{-}118)^*$$

Equations (9-114) and (9-118) are seen to consist of the mid-frequency gain of Eq. (9-112) multiplied by a factor that is a function of the frequency and with a phase angle equal to the 180-deg phase reversal at mid-frequency minus or plus an additional phase shift. If desired, Eqs. (9-112), (9-114), and (9-118) may be combined into a single equation as was done in Eq. (9-72), page 290.

In using these equations it must be recalled that the h-f factor applies only when the assumption is valid that R_L is sufficiently small to make the effect of the transformer and tube capacitances, C_T, negligible. Since this is the normal condition, the equation is useful, but care should be taken not to extend its application to special cases where it does not apply.

It is of interest to note that the maximum phase shift from mid-frequency to very high frequencies is 90 deg, not 180 deg as in Eq. (9-68), and that there is no indication of a resonant rise in voltage as in the latter equation. This is due to the virtual elimination of C_T by the shunting effect of R_L, this amplifier being an exaggerated case of the resistance-loaded amplifier described on page 302.

3. FEED-BACK AMPLIFIERS[1]

When a large fraction of the output voltage of an amplifier is fed back into the input circuit in such a phase as to oppose the impressed signal, the gain of the amplifier becomes substantially independent of μ, r_p, and g_m and of the direct electrode voltages, the frequency-response range is considerably increased at both the lower and the upper ends, amplitude distortion is materially reduced, and the effects of tube noise and hum voltages are decreased. Thus most high-quality amplifiers today are built with *negative* or *reversed feedback*.

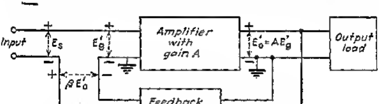


FIG. 9-68. Illustrating the principle of feed-back amplifier operation.

Analysis of Feed-back Amplifiers. The feed-back principle is illustrated in Fig. 9-68. The input of an amplifier having a gain **A** is supplied with a signal voltage $\mathbf{E}_s$ and with a portion $\beta\mathbf{E}_0'$ of its own output voltage, where the (′) mark in $\mathbf{E}_0'$ denotes output voltage under feed-back conditions as distinguished from the no-feed-back voltage $\mathbf{E}_0$. Note that β is a complex number indicating

[1] For a more detailed analysis than is presented here, see H. S. Black, Stabilized Feedback Amplifiers, *Bell System Tech. J.*, **13**, p. 1, January, 1934; or *Elec. Eng.*, **53**, p. 114, January, 1934. Also see H. Nyquist, Regeneration Theory, *Bell System Tech. J.*, **11**, p. 126, January, 1932.

that the feed-back circuit may alter the phase as well as the magnitude of the voltage fed back.

The no-feed-back gain **A** of an amplifier is equal to the ratio of the voltage appearing across the output terminals of the amplifier to the voltage impressed across the input terminals not including the feed-back circuit. From Fig. 9-68 this may be written

$$\mathbf{A} = \frac{\mathbf{E}_o'}{\mathbf{E}_g'} \qquad (9\text{-}119)$$

Of more interest is the ratio of output voltage to *total* input voltage which we will define as the gain of the amplifier *with* feedback, **A'**. From Fig. 9-68 the total input voltage may be written $\mathbf{E}_g = \mathbf{E}_g' - \beta\mathbf{E}_o'$; therefore, the over-all gain of the amplifier with feedback is

$$\mathbf{A}' = \frac{\mathbf{E}_o'}{\mathbf{E}_g} = \frac{\mathbf{E}_o'}{\mathbf{E}_g' - \beta\mathbf{E}_o'} \qquad (9\text{-}120)$$

If both numerator and denominator of Eq. (9-120) are divided by $\mathbf{E}_g'$, and $\mathbf{E}_o'/\mathbf{E}_g'$ is replaced by **A** from Eq. (9-119), the result is

$$\mathbf{A}' = \frac{\mathbf{A}}{1 - \mathbf{A}\beta} \qquad (9\text{-}121)^*$$

For the sake of simplicity let us consider a single-stage, resistance-coupled amplifier operating in the mid-frequency region, and a feed-back network that is purely resistive. Under these conditions **A** will be a negative number and β a positive one, the phase angle of each being zero. The denominator of Eq. (9-121) will, therefore, be larger than unity regardless of how large or how small β may be; therefore, the gain with feedback is less than the gain without feedback. Amplifiers exhibiting this characteristic are commonly known as *negative feed-back amplifiers*.

A valuable feature of the negative feed-back amplifier is that its gain may be made virtually independent of variations in tube characteristics or supply voltage. This may be seen from Eq. (9-121) by noting that if $\mathbf{A}\beta \gg 1$, **A'** may be written

$$\mathbf{A}' = -\frac{1}{\beta} \qquad \text{(approx)} \qquad (9\text{-}122)^*$$

being independent of the no-feed-back gain **A**. Thus the gain with feedback is independent of all factors that affect **A**, such as varia-

tions in supply voltages and of changes in tube coefficients due to tube replacements. This is a very valuable feature in laboratory instruments where permanency of calibration is desirable, or in multistage amplifiers as, for example, in long-distance telephone circuits where even a small change in the gain of each of several amplifier tubes would produce an excessive change in over-all gain.

Feedback also materially reduces amplitude distortion and the effect of noise and hum voltages in the output of an amplifier.[1] Since distortion is a function of the output voltage, we may write for the amplifier without feedback

$$\mathbf{D} = f(\mathbf{E}_0) \tag{9-123}$$

where D is the distortion voltage in the plate circuit with no feedback.[2]

Under feed-back conditions the total distortion voltage appearing in the plate circuit will be the sum of the distortion voltage introduced by the tube and that which has been fed back and amplified, or

$$\mathbf{D}' = f(\mathbf{E}_0') + \mathbf{A}\beta\mathbf{D}' \tag{9-124}$$

Solving this equation for $\mathbf{D}'$ gives

$$\mathbf{D}' = \frac{f(\mathbf{E}_0')}{1 - \mathbf{A}\beta} \tag{9-125}$$

Under the conditions that the output voltage of the amplifier with feedback is the same as the output without feedback (which means that the signal voltage on the grid is much higher with feedback than without) we may write

$$\frac{\mathbf{D}'}{\mathbf{D}} = \frac{1}{1 - \mathbf{A}\beta} \tag{9-126}^*$$

which shows that, *for the same output*, feedback will reduce the distortion by an amount dependent on the feed-back factor β.

[1] See also F. E. Terman, Feedback Amplifier Design, *Electronics*, **10**, p. 12, January, 1937; and Geoffrey Builder, The Effect of Negative Feedback on Power Supply Hum in Audio-frequency Amplifiers, *Proc. IRE*, **34**, pp. 140W–144W, March, 1946.

[2] D may also include voltages due to hum or tube noise introduced in the tube under consideration.

The increase in signal voltage required with feedback, to maintain the same output voltage as without feedback, may be found by first rewriting Eq. (9-121) as

$$\frac{A'}{A} = \frac{1}{1 - A\beta} \tag{9-127}$$

If the output voltage is to be the same with or without feedback, then $E_0' = E_0$ or $A'E_s' = AE_s$, which may be written $E_s'/E_s = A/A'$. Substituting the relation of Eq. (9-127), we find

$$\frac{E_s'}{E_s} = 1 - A\beta \tag{9-128}^*$$

or the signal voltage must be increased by a factor equal to the magnitude of $1 - A\beta$.

In designing a high-fidelity amplifier the cost of providing this increase in excitation required with feedback is small compared with the improvement in quality obtained, since additional stages of voltage amplification are easily added to provide the necessary driving voltage.[1]

Effect of Feedback on the Frequency Response of an Amplifier. The improvement in frequency response that feedback will effect is evident from Eqs. (9-121) and (9-122), since frequency distortion is due to a variation, with frequency, of the no-feed-back gain A. If Eq. (9-122) is even approximately true, it would appear that an amplifier should give substantially linear response no matter what type of load impedance might be used or what frequency might be impressed.

Unfortunately two things happen as the frequency approaches either the upper or the lower limit of the mid-frequency range. In the first place A becomes smaller, approaching zero, so that the approximation of Eq. (9-122) can no longer be used; second, the phase angle of A becomes other than the 180 deg of the mid-frequency region. If this phase shift is sufficient for the angle of $A\beta$ to approach zero, the feedback will become *positive* and the amplifier will oscillate if the magnitude of $A\beta$ is still sufficiently large. This means that an a-c signal will be maintained in the amplifier independently of the applied signal, the frequency of which is a

[1] For some suggested design procedures, see Stewart Becker, The Stability Factor of Negative Feedback Amplifiers, *Proc. IRE*, 32, pp. 351–353, June, 1944.

function of the circuit constants. Under such circumstances the circuit is useless as an amplifier.

To ascertain as to whether or not oscillations may be set up in the amplifier, the in-phase and quadrature components of Aβ and its conjugate should be plotted in rectangular coordinates for all frequencies at which the gain of the amplifier is greater than zero. It may be shown that if this curve does not enclose the point 1,0 the amplifier will not oscillate.[1] This is illustrated in Fig. 9-69 where a possible curve has been plotted, the part lying above the X axis being Aβ and the part lying below the X axis being its conjugate. The amplifier to which this curve applies will not oscillate since the curve does not enclose the point 1,0. In general two such curves must be drawn, one for the l-f end and the other for the h-f end of the amplifier frequency spectrum.

Throughout the mid-frequency range of a resistance-coupled amplifier, A will have a phase angle of zero if there are an even number of stages, and 180 deg if there are an odd number of stages. Let it be assumed that the feed-back network is purely resistive and, therefore, that the phase angle for β is either zero or 180 deg, depending on the circuit. There will then be no quadrature component for Aβ, and Aβ may be represented by a magnitude with either a plus or a minus sign but with no phase angle. Referring to Fig. 9-69 it is obvious that to prevent oscillations under these assumed conditions either Aβ must be negative, *i.e.*, lie to the left along the X axis, or, if positive, be less than unity. Normally it is made negative by suitable design of the feed-back network.

Similar conclusions may be drawn directly from Eq. (9-121) for the simple case where Aβ has no quadrature component. Thus if Aβ is a negative number with zero phase angle, $(1 - \text{A}\beta)$ will be a positive number greater than one and the gain A$'$ (with feedback) will be less than A (without feedback). On the other hand if Aβ is a positive number A$'$ will be larger than A, and if Aβ is also numerically larger than one the denominator will be negative and the amplifier will oscillate.

[1] H. Nyquist, Regeneration Theory, *Bell System Tech. J.*, 11, p. 126, January, 1932. Nyquist gives this rule: "Plot plus and minus the imaginary part of Aβ against the real part for all frequencies from zero to infinity. If the point $1 + j0$ lies completely outside this curve, the system is stable (will not oscillate); if not, it is unstable (will oscillate)."

Analysis of the more general case, where $A\beta$ may have any phase angle whatsoever, is of course more difficult; but if we again assume the feed-back circuit to be purely resistive, we can determine the effect of the phase angle of **A** from the equations developed in the early sections of this chapter. Thus a single-stage resistance-coupled amplifier with adequate by-passing of screen-grid and cathode circuits has, as may be seen from Eq. (9-30), a maximum phase shift of 90 deg leading at low frequencies and 90 deg lagging at high frequencies, relative to the mid-frequency phase. A single-stage amplifier of this type could not oscillate since the phase shift from mid-frequency would have to be 180 deg if it is to combine with the 180-deg phase reversal due to the tube and give $A\beta$ a phase angle of zero, a necessary condition for oscillation. Even a two-stage amplifier will not oscillate since, although the maximum phase shift from mid-frequency is 180 deg, causing the phase angle of $A\beta$ to approach zero at high and low frequencies, the amplitude of $A\beta$ also approaches zero, giving a curve similar to that of Fig. 9-69 which does not enclose the point 1,0. Any resistance-coupled amplifier using more than two stages is evidently subject to oscillations unless special precautions are taken.

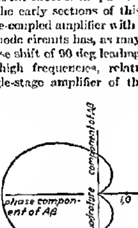


FIG. 9-69 Curve drawn to determine the possibility of oscillations in a amplifier.

Ineffective by-passing of the screen-grid circuit will increase the tendency of the amplifier to oscillate as may be seen from Eq. (9-48) where the total phase shift from mid-frequency approaches 180 deg leading at low frequencies in a single-stage amplifier. Satisfactory by-passing requires that the condenser C_d, Fig. 9-8b, be large enough to prevent the phase angle, $\tan^{-1}(X_d/r_{p2})$, becoming appreciably larger than zero until the gain of the amplifier has fallen low enough to decrease the magnitude of $A\beta$ nearly to unity.

Similar analysis may be made of the resistance-coupled amplifier with l-f and h-f compensation and of the transformer-coupled amplifier. In the latter, for example, the phase shift per stage approaches 180 deg at the h-f end of the spectrum [see Eq. (9-72)];

therefore, there is more tendency to oscillate than in a resistance-coupled amplifier. Power amplifiers using stepdown transformers in the output circuit, on the other hand, have a maximum phase shift of 90 deg from mid-frequency, not 180 deg [see Eq. (9-114)], except in the case of parallel feed (as in Fig. 9-21) where the phase shift approaches 180 deg at low frequencies owing to the blocking condenser.

When feedback is applied to three or more stages of resistance-coupled amplification (or two or more stages of any amplifier with a phase shift from mid-frequency of 180 deg per stage) special provisions must be incorporated to prevent oscillation. A possible method is to design one stage of a three-stage amplifier with a wider mid-frequency range, presumably at some sacrifice in gain. With such a design the two higher gain stages will produce a phase shift from mid-frequency approaching 180 deg at frequencies for which the phase shift in the lower gain stage will still be nearly zero. At frequencies sufficiently low or high to cause appreciable phase shift in the lower gain stage, the magnitude of the gain in the other two stages should have become so low that oscillation is impossible.[1]

Current Feedback. In the preceding analysis of feed-back amplifiers the feed-back voltage was proportional to the output voltage. It is also possible to design amplifiers in which the feed-back voltage is proportional to the output current. Figure 9-70 shows the principle of this type of operation, the feed-back voltage being set up across an impedance in series with the plate or output circuit.

The over-all gain of the circuit of Fig. 9-70, including feedback, is

$$\mathbf{A}' = \frac{\mathbf{E}_0'}{\mathbf{E}_s} = \frac{\mathbf{E}_0'}{\mathbf{E}_g' - \mathbf{I}_0'Z_f} \tag{9-129}$$

[1] For further information, see such references as F. E. Terman and Wen-Yuan Pan, Frequency Response Characteristic of Amplifiers Employing Negative Feedback, *Communications*, p. 5, March, 1939; Vincent Learned, Corrective Networks for Feedback Circuits, *Proc. IRE*, **32**, pp. 403–408, July, 1944; H. W. Bode, Relations between Attenuation and Phase in Feedback Amplifier Design, *Bell System Tech. J.*, **19**, p. 421, July, 1940; and H. W. Bode, "Network Analysis and Feedback Amplifier Design," D. Van Nostrand Company, Inc., New York, 1945.

If the impedance of the load is $\mathbf{Z}_L$, we may write $\mathbf{I}_0' = \mathbf{E}_0'/\mathbf{Z}_L$ and Eq. (9-129) reduces to

$$\mathbf{A}' = \frac{\mathbf{E}_0'}{\mathbf{E}_g' - \mathbf{E}_0' \dfrac{\mathbf{Z}_f}{\mathbf{Z}_L}} \tag{9-130}$$

If both numerator and denominator are divided by $\mathbf{E}_g'$ and the relation of Eq (9-119) substituted, the result is

$$\mathbf{A}' = \frac{\mathbf{A}}{1 - \mathbf{A} \dfrac{\mathbf{Z}_f}{\mathbf{Z}_L}} \tag{9-131*}$$

Since current feedback tends to maintain constant output *current* (for a given impressed voltage) regardless of changes in amplifier gain or load impedance, rather than constant output *voltage*, it

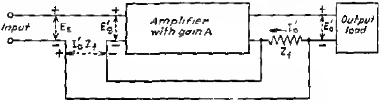
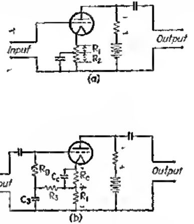


FIG 9-70 Simple current-feedback circuit

will not correct for frequency distortion due to changes in load impedance with frequency. Thus current feedback is less desirable than voltage feedback if the objective is constant over-all gain with a minimum of distortion

Feed-back Circuits. There are many types of feed-back circuits, and only a few will be presented here. In the current-feedback circuit of Fig 9-71*a* the same resistor is used both to provide grid bias and to set up a feed-back voltage. The alternating component of plate current must flow through the resistor R_1 and will set up a voltage that, with a resistive load in the plate circuit, is 180 deg out of phase with the input signal, in the mid-frequency range. Connecting a condenser across part of the resistance, R_2, reduces the feedback without affecting the bias, since only the resistance R_1 is then effective in producing a feed-back voltage, whereas the direct component of plate current must flow through R_2 as well, giving a bias voltage equal to $I_b(R_1 + R_2)$.

A modification of this circuit is shown in Fig. 9-71*b* where the functions of feedback and of cathode resistor are separated. Here only the direct voltage across R_c is applied to the grid, and the alternating voltage appears across R_1, since C_c by-passes the alternating component of plate current around R_c. The resistance of



Current-feedback

FIG. 9-71. Two feed-back circuits in which the feedback is proportional to the output current.

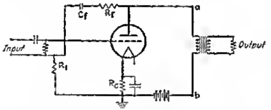


FIG. 9-72. A simple type of feed-back circuit wherein the feedback is a function of the output voltage.

R_g should be high compared with the impedance of the input circuit; that of R_3 should be large compared with R_1.

A simple voltage-feedback circuit is shown in Fig. 9-72. The feedback takes place from the plate through resistor R_f, setting up an emf across the resistor R_1. With resistance load as shown, the phase of the alternating plate voltage is exactly opposed to that of the grid voltage, in the mid-frequency range; therefore, the

emf set up across R_1 will be opposite in phase to the signal from the input circuit. The amount of feedback is controlled by the resistors R_f and R_1, and condenser C_f blocks off the direct plate voltage from the grid circuit. Since the feedback is proportional to the plate voltage, this circuit gives reductions in both amplitude and frequency distortion

The feed-back circuit of Fig. 9-72 may be applied to a push-pull amplifier, Fig. 9-73. The resistances R_1, R_1' correspond to the

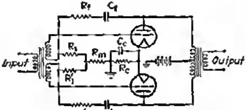


FIG. 9-73 Circuit of a push-pull amplifier with feedback.

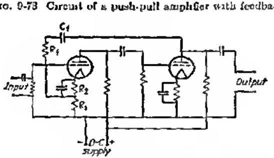


FIG 9-74. Feedback applied to two stages of a resistance-coupled amplifier.

resistance R_1 of Fig. 9-72, R_m is an additional resistance to improve the balance between the two tubes. Any unbalance between the tubes will produce a current through this resistance, thereby setting up a voltage of such polarity as to tend to restore the balance. The same arrangement may also be used with a phase-inverting tube to maintain balance.

In applying feedback to a two-stage amplifier, care must be exercised to reverse the polarity of the feedback from that obtained in the circuits of Figs. 9-71 and 9-72 to compensate for the additional phase reversal inserted by the added stage of amplification. Figure 9-74 shows the application of the circuit of Fig 9-72 to a

two-stage, resistance-coupled amplifier. The resistance R_1 sets up the feed-back voltage which is seen to be opposite in phase, with respect to the grid, from that set up by R_1 in Fig. 9-72. The resistance R_2 together with R_1 provides grid bias. Feedback is also provided in this circuit by plate current in the first tube flowing through R_1 (as in the circuits of Fig. 9-71) so that the total feedback is equal to the sum of the voltage set up by this current feedback and that due to the voltage feedback through R_f. Usually the current feedback is small compared to the voltage feedback.

Feedback may also be used with pentode amplifiers. When pentodes are used in the circuits of Figs. 9-71 and 9-74, it is necessary to return the screen-grid bypass condenser and the suppressor grid directly to the cathode rather than to ground. In addition, the resistance in series with the d-c supply to the screen must be sufficiently high to prevent the screen by-pass condenser from shunting out the feed-back resistor R_1. This is illustrated in Fig. 9-75 (similar to Fig. 9-71*a* for the triode) where R_d must be large compared with R_1.

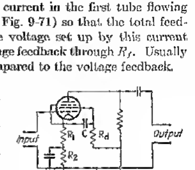


FIG. 9-75. Pentode amplifier with feedback.

Cathode-follower Amplifier. The circuit of a cathode-follower amplifier is shown in Fig. 9-76. It is essentially that of a resistance-coupled amplifier with 100 per cent feedback ($\beta = 1.0$); and the gain is therefore found from Eq. (9-121) to be

$$\mathbf{A}' = \frac{\mathbf{A}}{1 - \mathbf{A}} \tag{9-132}$$

As **A** is made larger without limit, the cathode-follower amplifier will evidently produce a maximum gain of one. It is, therefore, not used to provide voltage amplification but is an excellent power amplifier and impedance-transforming device.

The impedance-transforming characteristics of the cathode-follower amplifier may be demonstrated by expanding Eq. (9-132). First we must substitute for **A** from Eq. (9-119) whence (9-132) becomes

$$\mathbf{A}' = \frac{\mathbf{E}_0'}{\mathbf{E}_g' - \mathbf{E}_0'} \tag{9-133}$$

But

$$\mathbf{E}_o' = -\mathbf{E}_L = \mathbf{I}_p' Z_L \tag{9-134}$$

where $\mathbf{Z}_L$ is the parallel impedance of Z_o and the output circuit (Fig 9-76). $\mathbf{E}_g'$ may be found from Eq. (3-23)

$$\mathbf{E}_g' = -\mathbf{I}_p' \frac{r_p + Z_L}{\mu} \tag{9-135}$$

Substituting Eqs. (9-134) and (9-135) into (9-133) gives

$$\mathbf{A}' = \frac{-\mu Z_L}{r_p + Z_L + \mu Z_L} \tag{9-136}$$

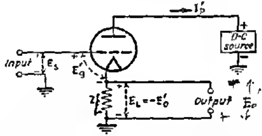


Fig. 9-76 Circuit of a cathode-follower amplifier.

If both numerator and denominator of this equation are divided by $(1 + \mu)$ the result is

$$\mathbf{A}' = \frac{\mathbf{E}_0'}{\mathbf{E}_s} = \frac{-\frac{\mu}{1+\mu} Z_L}{\frac{r_p}{1+\mu} + Z_L} \tag{9-137}$$

This equation should be compared with Eq. (9-119) for A which may be expanded as

$$\mathbf{A} = \frac{\mathbf{E}_0'}{\mathbf{E}_g'} = \frac{-\mu Z_L}{r_p + Z_L} \tag{9-138}$$

Equation (9-137) shows that the cathode-follower amplifier is the equivalent of a tube with amplification factor of $\mu/(1 + \mu)$ and a plate resistance of $r_p/(1 + \mu)$. If a tube with a high μ is used, such as a pentode or beam tube, the effective plate resistance is

very much smaller than that of the actual tube, being approximately $r_p/\mu = 1/g_m$.

It should now be evident that the equivalent circuit of the cathode-follower amplifier is as shown in Fig. 9-77. Theoretically maximum power output will be obtained when the output load is a pure resistance equal to the equivalent plate resistance $r_p/(1 + \mu)$. Since the equivalent plate resistance is very low, this type of amplifier can be connected to a very low load resistance without the use of an output transformer. When so used, the gain A of the amplifier without feedback becomes small, and many of the advantages of feed-back amplifiers, such as reduction of distortion, are partially or wholly lost. However the h-f response of the amplifier is excellent since R_o in Eq. (9-30) is very low in this amplifier. The low value of R_0 is due both to a low equivalent r_p and to a low R_1, as shown by Eq. (9-10). Amplitude distortion may be determined by applying methods similar to those used for conventional amplifiers.[1]

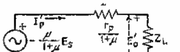


FIG. 9-77. Equivalent circuit of a cathode-follower amplifier.

The input impedance of a cathode follower is higher than that of conventional amplifiers. This may be seen from the equivalent circuit of Fig. 9-78 showing the capacitances only, corresponding to that of Fig. 9-2c. The capacitance C_{gk} is the grid-cathode capacitance of a triode and is essentially the control-grid–cathode plus the control-grid–screen-grid capacitance of a pentode or beam tube. The capacitance C_{gp} is that which exists between the plate and the control grid.

Following the method of page 261 we may write $\mathbf{E}_0' = -A'\mathbf{E}_s\underline{/\psi}$ where A' is the magnitude only of the gain of the cathode-follower amplifier. The following equations may now be written directly from the circuit of Fig. 9-78.

$$\mathbf{I}_g = \mathbf{I}_g' + \mathbf{I}_g'' \tag{9-139}$$

$$\mathbf{I}_g' = j\omega C_{gk}(\mathbf{E}_s + \mathbf{E}_0') = j\omega C_{gk}(\mathbf{E}_s - A'\mathbf{E}_s\underline{/\psi}) \tag{9-140}$$

$$\mathbf{I}_g'' = j\omega C_{gp}\mathbf{E}_s \tag{9-141}$$

[1] Graphical solutions designed especially for cathode-follower amplifiers are given by William A. Huber, Graphical Analysis of Cathode-biased Degenerative Amplifier, *Proc. IRE*, **35**, p. 265, March, 1947; and David L. Shapiro, The Graphical Design of Cathode-output Amplifiers, *Proc. IRE*, **32**, pp. 263–268, May, 1944.

Equation (9-140) may be rewritten in rectangular coordinates as

$$\mathbf{I}_g' = j\omega C_{gk} [\mathbf{E}_s - A'\mathbf{E}_s(\cos\psi + j\sin\psi)] \qquad (9\text{-}142)$$

We may now solve for the input admittance

$$Y = \mathbf{I}_g/\mathbf{E}_s = A'\omega C_{gk}\sin\psi + j\omega(C_{gp} + C_{gk} - A'C_{gk}\cos\psi) \qquad (9\text{-}143)^*$$

Equation (9-143) should be compared with Eq. (9-7) for a conventional grounded-cathode amplifier. The term in the parentheses is evidently the equivalent input capacitance C_g

$$C_g = C_{gp} + C_{gk}(1 - A'\cos\psi) \qquad (9\text{-}144)^*$$

and it is evident from this equation that this capacitance is even smaller than the geometric capacitance of the tube ($C_{gp} + C_{gk}$).

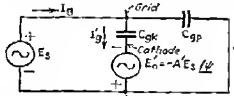


FIG. 9-78. Equivalent circuit of the input capacitance of a cathode follower amplifier.

This low input capacitance materially improves the performance of the driving amplifier. Decreased input capacitance is characteristic of all negative-feedback amplifiers, but the maximum decrease occurs in the cathode-follower amplifier.

The first term of Eq. (9-143) is of the nature of a conductance and represents feedback due to the output voltage E_0 acting through the grid-cathode capacitance. At audio frequencies this effect is negligible.[1]

The impedance Z_c in the circuit of Fig. 9-76 is commonly a resistance that provides grid bias as well as serving as part of the load, an arrangement that may be objectionable since a suitable value of resistance for load may not provide the correct bias.

[1] For information on the input admittance of this amplifier and of other similar circuits, see H. J. Reich, Input Admittance of Cathode-follower Amplifiers, *Proc. IRE*, 35, p. 573, June, 1947.

Figure 9-79 shows one method for avoiding this difficulty by providing a means of adjusting the bias independently of the load. The resistance ($R_1 + R_2$) is considerably larger than the impedance of the load so that it has little effect on the total output impedance of the amplifier. By adjusting the relative size of these two resistors the bias voltage may be made any value less than the total direct voltage appearing across the load impedance.

In actual practice the load impedance may consist of a transmission line, suitably terminated. If the transmission-line impedance is too high, a resistance may be bridged across the sending end of the line; if the line impedance is too low, a resistance may be inserted in series. The plate current of the amplifier must normally flow through the line and its terminating impedance, but if a resistance is connected in parallel with the sending end of the line, the direct component of the plate current may be confined to

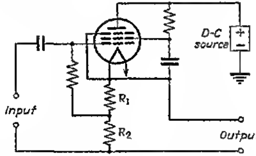


FIG. 9-79. Cathode-follower circuit in which the grid bias is adjustable independently of the load.

this resistance by inserting a suitably large blocking condenser in series with the line.[1]

Regeneration. Feedback may take place in amplifiers even when no special feed-back circuits are provided, owing to the presence of impedances common to both input and output circuits of an amplifier. When the feed-back signal is of such phase as to add to the desired signal (positive feedback), this phenomenon is commonly known as *regeneration*; when it is subtractive (negative feedback), the phenomenon is known as *degeneration*.

[1] Additional circuits and applications of the cathode-follower amplifier, at both audio and radio frequencies, are given by Kurt Schlesinger, Cathode-follower Circuits, *Proc. IRE*, **33**, pp. 843–855, December, 1945.

Regeneration may cause a number of very undesirable effects If of sufficient magnitude and proper phase, it may cause the amplifier to oscillate at an audible resonant frequency and so produce a howl in a loud-speaker connected to its output. In resistance-coupled amplifiers it may produce motorboating, so named because of the "put-put" sound frequently produced in the loud-speaker. Motorboating is caused by enough energy being fed back to build up a charge on the coupling condenser sufficient to cut off the flow of plate current through the tube, after which the condenser discharges through the grid leak at a rate depending on the size of the condenser and grid leak [1] The tendency to motorboat increases as the condenser and leak are increased in size and, therefore, as the l-f response of the amplifier is improved. Thus it is especially necessary to avoid regeneration in resistance-coupled amplifiers with good l-f response.

Regeneration may be caused by coupling through inductive fields, as between the windings of transformers Such sources of feedback may be eliminated by the use of resistance coupling or by thoroughly shielding the transformers and by so locating them in the amplifier as to prevent the stray field of one from linking with another. Capacitive coupling between wiring and tubes may also cause trouble, especially at the higher frequencies. Such capacitances are not likely to be of sufficient size to cause much regeneration at audio frequencies, and care in locating the wiring will usually avoid any trouble.

Impedances in the common return path from two or more tubes probably cause most of the trouble experienced with regeneration in amplifiers. Figure 9-80 shows the circuit of a three-stage, resistance-coupled amplifier with a single power supply for all tubes. The impedance between the output terminals AB of the rectifier is not zero even though shunted with a very large condenser Thus the alternating plate current of the third tube will set up a small voltage between the points A and B which will in turn be impressed on the plate of the first tube and therefore on the grid of the second tube. This voltage is of proper phase to increase the alternating plate current in the third tube and thus produce regeneration

[1] This action is similar to the performance of multivibrators (p 481) where feedback (or regeneration) is intentionally introduced

Resistance-capacitance filters are often used as "decoupling" circuits to prevent regeneration. Figure 9-81 shows the method of applying such filters to the amplifier of Fig. 9-80 to reduce the

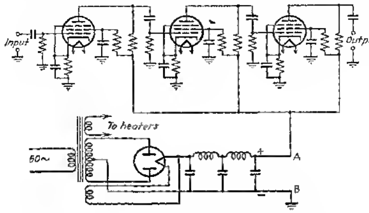


FIG. 9-80. Circuit of a resistance-coupled amplifier with common impedance in the plate supply, between points A and B.

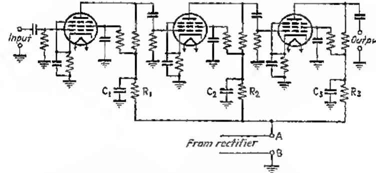


FIG. 9-81. Showing method of inserting resistance-capacitance filters in the circuit of Fig. 9-80, to reduce regeneration.

effect of the common rectifier impedance AB, the filters consisting of the resistances R_1, R_2, R_3 and the capacitances C_1, C_2, C_3. The alternating plate current of the third tube will now divide between the capacitance C_3 and the circuit consisting of the resistance R_3 and the rectifier in series. It is a relatively simple matter to make

the resistance of R_3 much larger than the reactance of C_3 and thus reduce the voltage built up between points A and B to a negligible quantity. The small voltage that is yet developed between A and B is further reduced by the action of R_1 and C_1 before any regenerative signal can reach the plate of the first tube.

The filter for the second tube, R_2C_2, prevents *degeneration*, since the phase of any voltage set up in the plate circuit of the second tube is such as to *decrease* alternating plate current in the third tube.

It is interesting to note that the decoupling circuit of Fig. 9-81 is the same as the l-f compensating circuit of Fig. 9-12. Thus l-f compensation automatically provides decoupling at higher frequencies.

Problems

9-1. The input to a certain amplifier is 40 mv and the output is 10 volts. The input impedance is 500 ohms (resistive), and the output impedance is 10 ohms (resistive). (*a*) What is the gain of the amplifier in db? (*b*) What is the output of the amplifier in dbv?

9-2. A certain triode has internal capacitances as follows: $C_{in} = 4.2$ $\mu\mu f$, $C_{out} = 2.8$ $\mu\mu f$, $C_{gp} = 2.6$ $\mu\mu f$. The tube is used as a resistance-coupled amplifier with a voltage gain of 60 (35.6 db). The phase shift is zero at the frequency under consideration. What is the equivalent capacitance presented by this tube to the preceding (or driving) amplifier?

9-3. Repeat Prob. 9-2 for a pentode with a gain of 100 (44.1 db) under similar operating conditions. The internal capacitances are $C_{in} = 5$ $\mu\mu f$, $C_{out} = 8$ $\mu\mu f$, $C_{gp} = 0.005$ $\mu\mu f$.

9-4. The constants of a resistance-coupled, triode amplifier, as in Fig. 9-4, are $\mu = 100$, $r_p = 90{,}000$ ohms, $R_1 = 0.25$ megohm, $R_g = 0.5$ megohm, $C = 0.005$ μf. What is the mid-frequency voltage gain? (The values of μ and r_p to be used in resistance-coupled amplifiers—especially r_p—usually differ from the values given in tube manuals, since the plate voltage is usually low in a resistance-coupled amplifier owing to the drop through the coupling resistance R_1.)

9-5. The capacitances of the tubes in the amplifier of Prob. 9-4 are as given in Prob. 9-2, and the capacitances of the wiring and other external elements between stages is an additional 7 $\mu\mu f$. Assume the gain of the driven tube to be equal to the mid-frequency gain computed from Prob. 9-4 and its phase shift to be zero (this gives a band width less than actually realizable). At what maximum frequency is the gain of the amplifier down 3 db from the mid-frequency gain?

9-6. At what minimum frequency is the gain of the amplifier of Probs. 9-4 and 9-5 down 3 db from the mid-frequency gain?

9-7. The triode tubes of Prob. 9-4 are replaced by pentode tubes in which $g_m = 1500$ μmhos, $r_p = 1.5$ megohms, and the capacitances are as given in Prob. 9-3 with wiring capacitance between tubes introducing an additional 7 $\mu\mu f$ in shunt with R_1. (*a*) What is the mid-frequency gain? (*b*) At

what maximum and minimum frequencies is the gain down 3 db from the mid-frequency gain? (Assume the screen-grid and cathode by-pass condensers of the driving tube to be of such size that they have negligible effect.)

9-8. (*a*) What must be the value of the coupling resistance R_1 for the amplifier of Prob. 9-7 if the gain of the amplifier is to remain constant to within 3 db of the mid-frequency gain up to a frequency of 2 Mc? (All other factors unchanged.) (*b*) What is the mid-frequency gain with this value of coupling resistor? (*c*) What is the minimum frequency at which the gain is down 3 db with this value of coupling resistance?

9-9. What is the db drop in gain at the minimum frequency computed in Prob. 9-7 if the effect of the screen-grid by-pass condenser is taken into account. $C_d = 0.2\ \mu f$, $r_{g2} = 20{,}000$ ohms, $R_d = 1$ megohm.

9-10. A three-stage, resistance-coupled amplifier is to be built, each stage having the constants of Prob. 9-7. Determine (*a*) the mid-frequency gain, (*b*) the maximum and minimum frequencies at which the gain of this amplifier is 3 db down from the mid-frequency gain.

9-11. The circuit of Fig. 9-12 is to be used with the amplifier of Prob. 9-8 to improve the l-f response. (*a*) Compute the value of C_2 to reduce the loss in gain from 3 to 0.5 db at the minimum frequency computed in Prob. 9-8. (*b*) Compute the minimum value of R_2 assuming that the terms on the two sides of the inequalities listed in the paragraph preceding Eq. (9-56) are in a ratio of not less than 50:1.

9-12. The circuit of Fig. 9-14 is to be used with the amplifier of Prob. 9-8 to improve the h-f response. Compute the inductance of L_1 to reduce the loss in gain from 3 to 0.5 db at 2 Mc.

9-13. A certain transformer-coupled, voltage amplifier has the following constants (see the equivalent circuit of Fig. 9-19*b*): $\mu = 20$, $r_p = 15{,}000$ ohms, $L_p = 150$ henrys, $L_s = 0.3$ henry, $R_s = 220$ ohms, $C_T = 60\mu\mu f$, $n = 2.8$. Compute (*a*) the mid-frequency gain, (*b*) the minimum frequency for a loss in gain of 3 db over the mid-frequency gain, (*c*) the frequency of maximum gain and the db rise in gain at this frequency above the mid-frequency gain.

9-14. Compute the power output and percentage distortion for a triode amplifier tube having the characteristics of Fig. 9-45. Operating point at $E_b = 275$, $E_c = -56$. Assume $E_{gm} = 56$ volts, $R_L = 3000$ ohms. (*b*) Repeat (*a*) for $R_L = 10{,}000$ ohms. (*c*) Compute the plate loss at full output under the conditions of part (*b*). (*d*) Repeat (*c*) but with no a-c grid voltage (zero output).

9-15. Compute the power output and percentage distortion for a pentode amplifier tube having the characteristics of Fig. 9-49. Operating point at $E_b = 250$, $E_c = -16.5$. Assume $E_{gm} = 16.5$ volts, $R_L = 2000$ ohms. (*b*) Repeat (*a*) for $R_L = 5000$ ohms. (*c*) Compute the plate loss at full output under the conditions of part (*a*). (*d*) Repeat (*c*) but with no a-c grid voltage (zero output).

9-16. Compute the plate efficiency (*a*) for the triode tube of Prob. 9-14 and (*b*) for the pentode tube of Prob. 9-15 making computations for both the given load resistances. (*c*) If the screen-grid current in the pentode is

10 ma at 250 volts, what is the efficiency of the pentode considering both screen-grid and plate losses?

9-17. Using the characteristic curves of Fig. 9-57, (a) compute the power output and distortion (third harmonic) for the load line shown ($E_b = 250$, $E_c = -50$, $E_{gm} = 50$). (b) Compute the plate efficiency.

HINT: For part (b) determine the direct current flowing in each tube by analyzing each as a single-tube amplifier using the method of pages 328 to 335. This will determine the d-c input to the amplifier. The currents I_{b0}, $i_{b\ max}$, $i_{b\ min}$, i_2, i_3 for one tube may be found from the *dotted* curves of Fig. 9-57, for the same plate and grid voltages as were used in finding these currents for the equivalent tube from the composite curves in part (a).

9-18. (a) Compute the power output and distortion for the class B amplifier to which the curves of Fig. 9-62 apply for the load line shown ($E_{gm} = E_b$). (b) Compute the plate efficiency. (See hint in Prob. 9-17.)

9-19. A resistance-coupled feed-back amplifier has the following constants: $g_m = 1200$ µmhos, $r_p = 1.2$ megohms, $R_1 = 0.2$ megohm, $R_g = 0.5$ megohm, $C = 0.01$ µf, $C_T = 22$ µµf. The feed-back factor β is the same at all frequencies for which the gain of the amplifier without feedback is within 6 db of the mid-frequency gain, being $0.1\angle 0$. (a) Compute the complex mid-frequency gain without feedback (A) and with feedback (A′). (b) Compute the minimum frequency at which the gain of the amplifier without feedback is down 3 db from the mid-frequency gain A. Then compute the complex gains A and A′ at this frequency. (c) Repeat (b) for the maximum frequency for a 3 db drop in gain without feedback. (d) How many db down from its mid-frequency value is the gain of the feed-back amplifier at the frequencies of parts (b) and (c)?

Design Problems[1]

9-20. Using the characteristic curves of Fig. 9-45, compute the power output, plate efficiency, and percentage distortion for various values of R_L for $E_b = 250$ volts, $E_c = -50$ volts, $E_{gm} = 20$ volts for the power output and efficiency and 50 volts for the distortion. Plot curves vs. R_L. These should be similar to the curves of Fig. 9-44.

9-21. Using the characteristic curves of Fig. 9-45, compute the power output, plate efficiency, and peak signal voltage for $E_b = 250$ volts, $E_c = E_{gm}$, letting E_{gm} and E_c so vary with R_L as to maintain a constant distortion of 5 per cent. Plot curves vs. R_L. These should be similar to the curves of Fig. 9-41.

9-22. The total primary inductance of a certain 3:1 ratio audio transformer is 150 henrys; the secondary inductance is 1350 henrys. Primary and secondary leakage reactances are each 0.2 per cent. Resistances of primary and secondary windings are 110 and 1000 ohms, respectively. The equivalent secondary capacitance including the tube capacitance (C_T in Fig. 9-19) is 60 µµf. If this transformer is used with a triode tube operating at $E_b = 250$ volts and $E_c = -8$ volts, where $\mu = 20$, $r_p = 10{,}000$ ohms, compute and plot the gain per stage in db vs. frequency, plotting frequency on a log

[1] See footnote on p. 215.

scale. Assume the input resistance of the driving and driven tubes to be the same. Vary the frequency from 25 to 25,000 cycles/sec.

9-23. Repeat Prob. 9-22 with a resistance of 100,000 ohms connected across the secondary terminals of the transformer.

9-24. Plot a set of composite characteristic curves for a push-pull, class A amplifier using the tube to which Fig. 9-15 applies, $E_b = 275$ volts, $E_c = -56$ volts, and determine the power output and plate efficiency for various load resistances. Plot curves vs. load resistance. (Assume $E_{gm} = E_c$ and that transformer coupling to the load is used with a zero-resistance primary winding. Also see the hint in Prob. 9-17.)

9-25. Compute the gain vs. frequency and phase angle vs. frequency curves of the amplifier of Prob. 9-19 assuming β remains constant to minimum and maximum frequencies for which the gain is down 3 db from the mid-frequency gain. Computations should cover such a frequency band as to include these two half-power points. Also plot $A\beta$ and its conjugate and see if this plot encloses the point 1,0.

CHAPTER 10

RADIO-FREQUENCY AMPLIFIERS

Vacuum tubes may be used to amplify r-f signals as well as those of audio frequency. However, it was shown in the preceding chapter that the shunting capacitance of the driven tube caused the amplification of audio amplifiers to approach zero at the higher frequencies. Therefore, if the vacuum tube is to be used as an amplifier at radio frequencies, some changes in circuits or tubes must be made. Actually most r-f amplifiers use a parallel resonant circuit as the load impedance, the input capacitance of the driven tube being included as a part of the resonating capacitance It then makes little difference as to how large this capacitance may be except that a large tube capacitance somewhat reduces the maximum frequency to which the amplifier may be tuned. The effect of the tube capacitance is further reduced by the nearly universal use of pentodes in low-power, r-f amplifiers, the input capacitance of which is less than that of corresponding triodes as was shown on page 264

The use of tuned circuits as load impedances in r-f amplifiers is made both possible and desirable by the type of signal to be amplified, most r-f signals containing only a single frequency component or at most a group of components spread over a rather narrow band of frequencies (as in a modulated wave). The resonant circuit presents a high impedance at this frequency (or band of frequencies), whereas its impedance is lower at all other frequencies, decreasing rapidly as the frequency shifts from the resonant value. The amplifier therefore has a high gain at resonance and a very low gain at any frequency that differs appreciably from resonance. It will therefore respond to radio signals of one frequency only, even though many different signals may be impressed on its input as in the case of an amplifier handling signals picked up by a radio antenna. This ability of the amplifier to select a desired r-f signal and reject all others is one of the important features of a well-designed r-f amplifier, since it makes pos-

sible the simultaneous transmission of many communications through space without interference.

Radio-frequency amplifiers, like those for a-f service, may be divided into two classes: voltage amplifiers and power amplifiers. The performance of a voltage amplifier is determined by its gain and its band width, which are the same characteristics as were studied for a-f amplifiers except that band width was treated under the heading of frequency distortion. Similarly, the performance of a power amplifier is determined largely by its ability to deliver as much power as possible with a minimum of distortion, although the distortion requirements are quite different at radio frequencies and at audio frequencies, as will be shown in later sections of this chapter. At radio frequencies voltage amplifiers are usually class A while power amplifiers are nearly always class B or C.

1. VOLTAGE AMPLIFIERS

Impedance-coupled Amplifiers. The simplest type of r-f voltage amplifier is that shown in Fig. 10-1*a* where the tuned circuit consists of an air-cored coil L with a tuning condenser C across its terminals. This condenser is generally of the variable, air-insulated type and may be varied throughout quite a wide range to tune the amplifier to the desired frequency. Iron is seldom used in coils at radio frequencies owing to skin effect[1] which renders it largely noneffective and to the iron losses which increase with frequency. Some work has been done with iron cores at the lower radio frequencies by using very finely powdered iron (or magnetic alloys of various types) held together with a suitable binder, but the results have not met with much favor.

The performance of the circuit of Fig. 10-1*a* may be determined by using the equivalent circuit of Fig. 10-1*b*. This follows the method used in analyzing resistance-coupled amplifiers in Chap. 9 where Norton's theorem was applied to obtain the constant-current generator delivering a current $\mathbf{I} = -g_m\mathbf{E}_g$.

[1] "Skin effect" refers to the tendency of flux to flow only along the surface of a piece of iron owing to the presence of eddy currents. Eddy currents may be thought of as the equivalent of a short-circuited secondary on a transformer; and if such a secondary is of low resistance, very little flux can flow through it. In the case at hand the induced emf in the eddy-current path becomes very large at high frequencies, causing a large eddy current to flow, and so materially reducing the flow of flux through the middle of each lamination.

The first step in determining the amplifier performance is to find the gain at resonance where the reactance of the condenser C is essentially equal to that of the coil L.[1] Let this reactance be

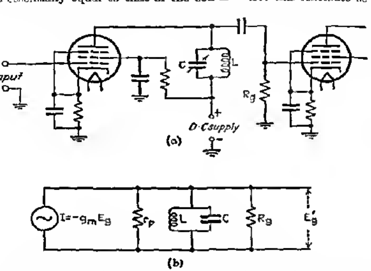


FIG. 10-1. Actual circuit (a) and equivalent circuit (b) of a tuned, r-f, impedance coupled amplifier

X. The resonant impedance of the parallel resonant circuit may then be found from

$$Z = \frac{1}{\frac{1}{-jX} + \frac{1}{R + jX}} = \frac{X^2}{R} - jX = XQ - jX \qquad (10\text{-}1)$$

where $Q = X/R$. In properly designed circuits Q is always much larger than 10 so that the second term in the last expression of Eq. (10-1) is negligible, and we may write

$$Z = XQ \qquad (10\text{-}2)$$

[1] In parallel resonant circuits the resonant frequency may be defined either as that at which the power factor is unity or that at which the impedance is a maximum. Each definition gives a slightly different frequency but, if the resistance of the coil is less than about 10 per cent of the reactance ($Q > 10$), the difference becomes negligible and the frequency is, for all practical purposes, that at which the two reactances are equal.

The output voltage $\mathbf{E}_o'$ of the amplifier of Fig. 10-1 is equal to the equivalent generator current $\mathbf{I}$ times the total impedance or, since $\mathbf{I} = -g_m\mathbf{E}_g$,

$$\mathbf{E}_o' = -g_m\mathbf{E}_g\left(\frac{1}{\frac{1}{XQ} + \frac{1}{r_p} + \frac{1}{R_g}}\right)$$

$$= -g_m\mathbf{E}_g X\frac{Q}{1 + \frac{XQ}{r_p} + \frac{XQ}{R_g}} \qquad (10\text{-}3)$$

By analogy with Eq. (9-15) for the resistance-coupled amplifier we should be able to write this equation as

$$\mathbf{E}_o' = -g_m\mathbf{E}_g\mathbf{Z}_0$$

where $\mathbf{Z}_0$ is the impedance of the entire load circuit at resonance. By analogy with Eq. (10-2) we may consider that $Z_0 = XQ'$ and Eq. (10-3) becomes

$$\mathbf{E}_o' = -g_m\mathbf{E}_g XQ' \qquad (10\text{-}4)$$

where Q' is the effective Q of the entire circuit and is found from Eq. (10-3) to be

$$Q' = \frac{Q}{1 + \frac{XQ}{r_p} + \frac{XQ}{R_g}} \qquad (10\text{-}5)$$

The gain of the amplifier at resonance may now be written from Eq. (10-4) as

$$\mathbf{A} = \frac{\mathbf{E}_o'}{\mathbf{E}_g} = g_m XQ'\underline{/180^\circ} \qquad (10\text{-}6)^*$$

The second step in determining the amplifier performance is to find the band width. We may define band width as the difference between the frequency f_2 above resonance at which the gain is 70.7 per cent of the maximum, and the frequency f_1 below resonance at which the gain is 70.7 per cent. These two frequencies are indicated in Fig. 10-2.

A simple procedure for finding $f_2 - f_1$ is to start with the following commonly used expression for Q:[1]

$$Q = \frac{f_r}{f_2 - f_1} \qquad (10\text{-}7)$$

[1] For derivation of this expression, see any book on resonant circuit theory such as W. L. Everitt, "Communication Engineering," pp. 63–65, 2d ed., McGraw-Hill Book Company, Inc., New York, 1937.

where f_r is the resonant frequency as indicated in Fig. 10-2. In the circuit under consideration the effective Q is Q' as given by Eq. (10-5). We may, therefore, replace Q with Q' in Eq. (10-7) and solve for the band width.

$$\text{Band width} = f_2 - f_1 = \frac{f_r}{Q'} \qquad (10\text{-}8)^*$$

Equations (10-6) and (10-8) provide sufficient information to design most r-f amplifiers. The desired band width is known and Q' may be computed from Eq (10-8). A tube is selected, which determines g_m. The only unknown remaining on the right-hand side of Eq. (10-6) is X, which should be made as large as

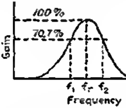


FIG. 10-2. Curve of the gain of a r-f amplifier showing the band width from f_1 to f_2

possible. The principal limitations in its magnitude are practical ones due to the physical dimensions of the coil and condenser, and distributed capacitances which may be excessive in large coils.

In designing the coil care should be taken that its resistance is sufficiently low as to give it a value of Q which, when substituted in Eq. (10-5), will give a value of Q' equal to that obtained from Eq (10-8). If this is found to be impractical, a different value of X must be selected until a coil with proper Q can be designed.

Transformer-coupled Amplifiers. A more commonly used amplifier is that of Fig 10-3. The coils L_1 and L_2 constitute an air-cored transformer which may provide some step-up action. The remaining circuit elements serve the same functions as in Fig. 10-1.

Analysis of this circuit may be expedited by using the equivalent circuit of Fig. 10-4, obtained in the same manner as was Fig. 10-1b. The capacitance C_t includes both the capacitance of the tuning condenser and the input capacitance of the next tube, together with any wiring and other capacitances present. R_1 is the re-

sistance of coil L_1. R_2 represents the equivalent resistance of the secondary and, therefore, includes the effect of the input resistance of the next (or driven) tube as well as the resistance of the coil L_2. In properly designed amplifiers the input resistance of the driven tube is so high that it may be neglected, and R_2 may normally be considered equal to the coil resistance alone without appreciable error.

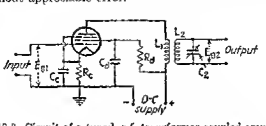


FIG. 10-3. Circuit of a tuned, r-f, transformer-coupled amplifier.

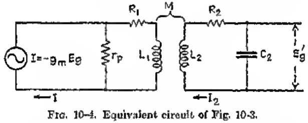


FIG. 10-4. Equivalent circuit of Fig. 10-3.

The impedance seen looking into the terminals of the coil L_1 is very much less than r_p; therefore, the current through L_1 may be assumed to be $\mathbf{I} = -g_m\mathbf{E}_g$ without serious error. The voltage induced into coil L_2 is therefore

$$\mathbf{E}_2 = -j\omega M\mathbf{I} = j\omega Mg_m\mathbf{E}_g \tag{10-9}$$

This voltage will cause a current $\mathbf{I}_2$ to flow in the secondary circuit. Since the impedance of the secondary at resonance is equal to the resistance alone, we may write

$$\mathbf{I}_2 = \frac{\mathbf{E}_2}{R_2} \tag{10-10}$$

The output voltage $\mathbf{E}_g'$ is equal to the voltage drop across the condenser C_2 or

$$\mathbf{E}_g' = \mathbf{I}_2\left(-j\frac{1}{\omega C_2}\right) = -j\frac{1}{\omega C_2 R_2}(j\omega M g_m \mathbf{E}_s) \qquad (10\text{-}11)$$

But at resonance $1/\omega C_2 = \omega L_2$ and Eq. (10-11) becomes

$$\mathbf{E}_g' = \frac{\omega L_2}{R_2}\omega M g_m \mathbf{E}_s \qquad (10\text{-}12)$$

Replacing $\omega L_2/R_2$ with Q, the gain may now be written

$$\mathbf{A} = \frac{\mathbf{E}_g'}{\mathbf{E}_g} = g_m\omega MQ \qquad (10\text{-}13)^*$$

The band width may be found in the same manner as for the impedance-coupled amplifier. In this case there is no resistance, other than that of the secondary coil, which is of such a magnitude as to be important and the effective Q is therefore equal to the actual Q of the secondary circuit. Therefore, we may write

$$\text{Band width} = \frac{f_r}{Q} \qquad (10\text{-}14)^*$$

where $Q = \omega L_2/R_2$

Equations (10-13) and (10-14) are sufficient to permit the design of amplifiers of this type with the information normally available.

Band-pass Filters. One of the most common applications of r-f amplifiers is in the amplification of modulated currents such as are encountered in radiotelephony. A modulated current contains many components having frequencies lying within a band of a definite width (see Chap. 13). In radiotelephony this band generally has a width of 10,000 to 20,000 cycles; and if a sharply resonant circuit is used in the amplifiers of Fig. 10-1 or 10-3, the amplification of certain frequency components will be greater than that of others. This results in distortion of the received signal (see Chap. 14).

In Fig. 10-5 are shown two resonance curves: 1 for a sharply resonant circuit and 2 for one less sharp, *i.e.*, containing more resistance and therefore having a lower Q. (The latter curve was drawn for a higher input voltage than the former, or its ordinates would be lower than those of 1 at all frequencies.) Evidently the circuit giving curve 1 will produce a higher gain than that for curve 2 but will not amplify all components in a 10,000-cycle band as well as will the latter. It is, therefore, necessary

to use a tuned circuit with a Q that is not too high, in spite of lower gain, to secure good quality of reproduced signal.

This sacrifice in gain to secure better quality of reproduction may be quite permissible, but a sacrifice in selectivity is not. By selectivity is meant the ability of the amplifier to reject currents from another undesired band (such as the one indicated by the

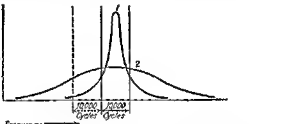


FIG. 10-5. Two possible resonance curves for the amplifiers of Figs. 10-1 and 10-3.

dotted lines). Evidently circuit 1 would amplify only a very small part of this undesired band, whereas circuit 2 would reproduce it with almost the same intensity as the desired band. This is a very serious problem in radio reception where two different stations may be allocated to immediately adjoining or adjacent bands (or channels).

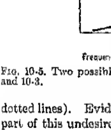


FIG. 10-6. Ideal shape of resonance curve for high fidelity and high selectivity.

The most desirable type of resonance curve would be one that is rectangular in shape with a width just equal to that of the band to be received, say 10,000 cycles (Fig. 10-6). Practically it is impossible to obtain this exact shape of resonance curve, but it may be approached rather closely by proper combinations of condensers and inductances. A very simple way to secure a bandpass effect is to use the regular transformer-coupled, r-f amplifier of Fig. 10-3 but place a tuning condenser across the primary as well as across the secondary (Fig. 10-7).

The output voltage of this circuit will be found to vary with frequency according to the curves of Fig. 10-8 where four curves are shown corresponding to four different values of the coupling

coefficient k, ($k = M/\sqrt{L_1L_2}$). It is assumed, for simplicity, that primary and secondary are tuned to the same frequency, *i e*, $L_1C_1 = L_2C_2$, and have the same Q, ($Q_1 = Q_2$), a condition commonly realized in practice.

At a low value of k the response curve is similar to that of a

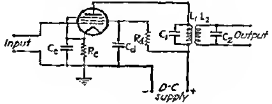


Fig 10 7 Radio-frequency amplifier with band-pass filter in the output circuit

single tuned circuit (curve k_1). Increasing k will at first result in an increase in the output voltage as in curve k_2. As k is further increased, a value is reached at which the output voltage is a maximum and continued increase in k will cause a reduction in the voltage at resonance. At the same time voltage peaks will appear on either side of the resonant frequency as shown in curves k_3 and k_4.

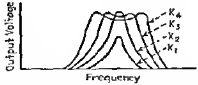


Fig 10-8 Frequency response curves of two coupled, tuned circuits for four different values of coupling.

The complete solution of the circuit of Fig 10-7 is rather difficult to obtain, but the gain at the resonant frequency may be found rather easily with the aid of the equivalent circuits of Fig 10-9. In Fig 10-9a the tube has been replaced by a constant-current generator through the use of Norton's theorem just as was done for the resistance-coupled amplifier in Chap 9 The plate resistance r_p is normally very much higher than the impedance seen

looking into the primary of the transformer and may be neglected. With this omission Thévenin's theorem may be applied at the points AB of Fig. 10-9a to give the circuit of Fig. 10-9b. The voltage of the equivalent generator would normally be written $(-g_m\mathbf{E}_g)(-j1/\omega C_1)$ but at resonance $1/\omega C_1 = \omega L_1$ and we may write the voltage as $jg_m\omega L_1\mathbf{E}_g$.

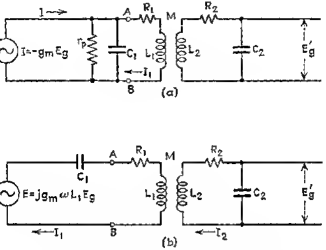


FIG. 10-9. Equivalent circuits for amplifier of Fig. 10-7.

Considering, for the moment, only the circuit to the right of the points AB, we may write

$$\mathbf{E}_2 = -j\omega M\mathbf{I}_1 \tag{10-15}$$

where $\mathbf{E}_2$ is the voltage induced in the coil L_2 by a current $\mathbf{I}_1$ flowing in the primary.

The induced voltage $\mathbf{E}_2$ will in turn cause a current $\mathbf{I}_2$ to flow in the secondary equal to this voltage divided by the secondary impedance. At resonance the impedance is merely the coil resistance R_2 and we may write

$$\mathbf{I}_2 = \frac{\mathbf{E}_2}{R_2} = \frac{-j\omega M\mathbf{I}_1}{R_2} \tag{10-16}$$

This current will in turn induce a voltage $\mathbf{E}_1$ back into the primary equal to

$$\mathbf{E}_1 = -j\omega M\mathbf{I}_2 = -\frac{(\omega M)^2}{R_2}\mathbf{I}_1 \qquad (10\text{-}17)$$

The current $\mathbf{I}_1$ flowing in the primary may now be written in terms of the voltages and impedances of that circuit

$$\mathbf{I}_1 = \frac{jg_m\omega L_1\mathbf{E}_g + \mathbf{E}_1}{R_1} = \frac{jg_m\omega L_1 E_g}{R_1} - \frac{(\omega M)^2}{R_1 R_2}\mathbf{I}_1 \qquad (10\text{-}18)$$

The reactances of the coil L_1 and the condenser C_1 do not appear in the denominator of this equation because they are equal in magnitude and opposite in sign and, therefore, cancel out.

Solving Eq. (10-18) for $\mathbf{I}_1$ gives

$$\mathbf{I}_1 = \frac{jg_m\omega L_1\mathbf{E}_g}{R_1 + \dfrac{(\omega M)^2}{R_2}} \qquad (10\text{-}19)$$

Equation (10-19) and Fig. 10-9*b* show that the primary current is equal to the voltage of the equivalent generator of Fig. 10-9*b* divided by the impedance of the primary circuit and that this impedance is equal to the resistance of the primary coil plus the reflected impedance from the secondary $(\omega M)^2/R_2$. Thus the secondary has the same effect on the primary as though an impedance of $(\omega M)^2/R_2$ were inserted in series with the primary with the secondary removed. [It may be shown that, if the secondary is not in resonance, the impedance reflected into the primary would be $(\omega M)^2/\mathbf{Z}_2$ where $\mathbf{Z}_2$ is the impedance of the secondary circuit.]

The gain of the amplifier may now be found by first determining the voltage $\mathbf{E}_g'$ set up across the condenser C_2. This voltage is given by

$$\mathbf{E}_g' = \mathbf{I}_2\frac{-j1}{\omega C_2} = \mathbf{I}_2(-j\omega L_2) \qquad (10\text{-}20)$$

The amplifier gain is, therefore,

$$\mathbf{A} = \frac{\mathbf{E}_g'}{\mathbf{E}_g} = \frac{-j\omega L_2\mathbf{I}_2}{\mathbf{E}_g} \qquad (10\text{-}21)$$

Substituting Eq. (10-19) into Eq. (10-10) and putting the resulting expression for I_2 into Eq. (10-21) gives

$$\mathbf{A} = \frac{-jg_m\omega L_1\omega MQ_2}{R_1 + \frac{(\omega M)^2}{R_2}} \tag{10-22}$$

where $Q_2 = \omega L_2/R_2$.

Equation (10-22) may be rewritten by substituting the relations $M = k\sqrt{L_1L_2}$ and $\omega L_1/R_1 = Q_1$ to give

$$\mathbf{A} = \frac{-jg_m\omega k\sqrt{L_1L_2}}{k^2 + (1/Q_1Q_2)} \tag{10-23)*$$

A study of Eq. (10-23) shows that the gain will first increase and then decrease as k is increased from zero toward its maximum possible value of 1. If the gain is differentiated with respect to k and the result set equal to zero to find the optimum value of k, it will be found that maximum gain will be obtained for $k = 1/\sqrt{Q_1Q_2}$. This value of k is known as the *critical value of coupling*, and any coupling greater than this critical value will result in a double-humped curve such as k_3 or k_4 of Fig. 10-8. If $Q_1 = Q_2 = Q$, as was assumed in plotting the curves of Fig. 10-8, the critical value of coupling will be that for which $k = 1/Q$. For band-pass operation the coupling coefficient must be made slightly greater than this critical value.

Development of the equations for the frequencies of the two peaks shown in Fig. 10-8 is more extensive than can be included here. However if k appreciably exceeds the critical value, the two peaks differ by a frequency of approximately kf_r where f_r is the frequency of resonance of primary and secondary circuits. The band width is, of course, slightly greater than the difference between these two peaks and is approximately equal to

$$\text{Band width} = 1.5\, kf_r \qquad \text{(approx)} \tag{10-24)*$$

Normal design would call for a value of k not too much larger than the critical value so as to secure a nearly flat response on either side of resonance, as in curve k_3 of Fig. 10-8.[1]

[1] For further design information, see Milton Dishal, Exact Design and Analysis of Double- and Triple-tuned Band-pass Amplifiers, *Proc. IRE*, 35, pp. 606–626, June, 1947.

Multistage Amplifiers. The band width of multistage amplifiers is always less than that of each stage. This must be so since, if the gain per stage at the highest and lowest frequencies is α times the mid-frequency gain, the gain at these frequencies for an n-stage amplifier will be α^n. Therefore, if the gain of the multistage amplifier is not to drop below 0.707 times the mid-frequency gain, at the upper and lower frequency limits, we may write

$$\alpha^n = 0.707 \tag{10-25}$$

from which we may solve for α to give

$$\alpha = (0.707)^{\frac{1}{n}} \tag{10-26}$$

This means that if all stages are identical, the upper and lower frequency limits of each stage must be at the frequency for which the gain is α times the mid-frequency gain, where α is given by Eq. (10-26). For most of the analysis thus far presented it was assumed that $\alpha = 0.707$ (i.e., $n = 1$).

2. POWER AMPLIFIERS

The variations in plate current in an r-f amplifier with resonant output circuit need not be confined to the straight portion of the characteristic curve of the tube as in the audio amplifier, since the output circuit is responsive to and reproduces only the original, or fundamental, component. Thus harmonics of the impressed radio frequency generated in the tube can set up no appreciable voltage in the plate circuit and are largely eliminated. This makes possible the use of amplifiers with a pulse type of plate current, operating at efficiencies as high as 85 per cent. Such high efficiency is of especial value in r-f power amplifiers where the output may run into hundreds or even thousands of kilowatts in high-power installations.

The reason for such high efficiency with a pulse type of plate current may be seen from Fig. 10-10. A sine-wave plate voltage is shown (curve a), but the plate current (curve b) consists of short pulses flowing only during *that portion of each cycle wherein the plate voltage is low*. Since the plate loss is equal to the average $e_b i_b$ and since i_b is zero at all times except when e_b is low, it is evident that the plate loss will be small. A larger output may be secured without sacrifice of efficiency by increasing the height but not the width of the current pulse. Since this will increase the input and

still prevent current flow at high values of e_b, the efficiency will remain high.

Amplifiers with pulselike plate current may be divided into two classes, designated as class B and class C.

A class B amplifier is defined (Appendix A) as "an amplifier in which the grid bias is approximately equal to the cutoff value so that the plate current is approximately zero when no exciting grid voltage is applied and so that plate current in a specific tube flows for approximately one-half of each cycle when an alternating grid voltage is applied." Such an amplifier produces a pulse of plate current of which both the peak and average values are practically proportional to the amplitude of the excitation voltage and is, therefore, termed a *linear amplifier*.

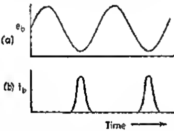


FIG. 10-10. Illustrating the possibilities of low plate loss with a pulse type of plate current.

A class C amplifier is defined (Appendix A) as "an amplifier in which the grid bias is appreciably greater than the cutoff value so that the plate current in each tube is zero when no alternating grid voltage is applied, and so that the plate current in a specific tube flows for appreciably less than one-half cycle when an alternating grid voltage is applied." Such an amplifier produces a power output proportional to the square of the plate voltage, within limits.

The wave shape of the plate current flowing in class B amplifiers is shown in Fig. 10-11. The current is seen to flow essentially only during the positive half cycle of the exciting voltage and to approximate a half sine wave in shape when a sinusoidal voltage is impressed. If it is assumed to be exactly a half sine wave and is analyzed by Fourier analysis,[1] the result is

[1] See Appendix C.

$$i_b = 0.318 I_m + 0.5 I_m \sin \omega t - 0.212 I_m \cos 2\omega t + \cdots \quad (10\text{-}27)$$

where I_m is the peak value of the plate-current pulse. Obviously the harmonic content in the output of this type of amplifier is very large

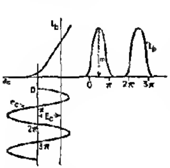


FIG. 10-11. Wave shape of the plate current in a class B amplifier

Figure 10-12 depicts the plate current flowing in a class C amplifier. It is seen to have a wave shape very like the top of one-half of a sine wave and if analyzed would be found to contain an even greater harmonic content than does that of the class B amplifier.

Field of Use of Class B and Class C Amplifiers. Class C amplifiers are used primarily in the amplification of unmodulated or of f-m[1] radio signals in radio transmitters or in other types of h-f equipment delivering appreciable power. Class B amplifiers are used in the amplification of a-m waves, as class C amplifiers are incapable of amplifying such waves without distortion (see page 430). Their efficiency is lower than that of class C amplifiers, and the latter are therefore preferred for all applications for which they are suited.

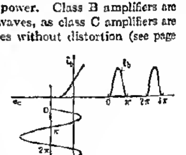


FIG. 10-12. Wave shape of the plate current in a class C amplifier

Voltage amplifiers may also be class B or C, if desired, rather than class A. However, such amplifiers find their principal applications in radio receivers where the power level of the amplified signal is so low that the efficiency is of no concern and where the cross modulation produced in class B or C amplifiers would seriously impair the performance of the receiver.[2] In radio transmitters

[1] See p. 508 for definitions of modulation.
[2] See p. 531 for a discussion of cross modulation.

and other high-power radio units, on the other hand, even an amplifier used to drive another tube must usually be designed along power amplifier lines, since in such service the grid of the driven tube normally swings quite positive on positive peaks of grid voltage and, therefore, draws very appreciable power.

Typical Circuits for Class B and Class C Amplifiers. Exactly the same circuits are used for both class B and class C amplifiers, the only distinction being in the amount of grid bias used. Figure

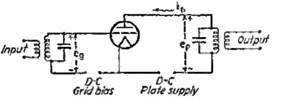


FIG. 10-13. Circuit of a class C amplifier using series feed. (Neutralizing circuits omitted.)

10-13 shows a typical circuit for triode tubes using series plate feed; *i.e.*, the direct component of plate current flows through the d-c supply and the load circuit in series.[1] A resonant circuit is used as the load and is tuned to the frequency of the signal im-

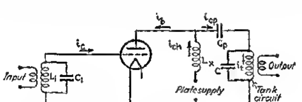


FIG. 10-14. Circuit of a class C amplifier using parallel feed. (Neutralizing circuits omitted.)

pressed on the grid of the tube. Grid bias is supplied by a fixed-voltage source (a method known as *fixed-bias*), although other methods are often used in practice (see page 422).

A more generally used form of class C amplifier circuit is shown in Fig. 10-14. It differs from the circuit of Fig. 10-13 principally

[1] A neutralizing circuit should be included when a triode is used, to avoid oscillations, but is omitted here and in subsequent diagrams to simplify the discussion. The need for neutralization is analyzed and typical circuits are presented in a later section (p. 423).

in that parallel feed is provided in the plate circuit, the direct component of plate current flowing through the r-f choke L_c, while the alternating component flows through the blocking condenser C_p. An advantage of this type of circuit is that one side of the tank[1] inductance is grounded to the cathode instead of the entire coil being above ground by the direct plate potential as in the series circuit (Fig. 10-13). Less insulation may therefore be used; and even if an arc to ground is initiated by the alternating voltage across the coil, no short-circuiting of the d-c supply will follow. The power supplies, not shown, usually consist of rectifiers drawing power from the 60-cycle mains.

Wave Shapes in Class C Amplifiers. Series Feed. Since a class B amplifier is, in a sense, a special case of a class C amplifier in which the bias is adjusted to approximately cutoff, the class C amplifier will be analyzed first. Figure 10-15 shows the wave shapes of the principal currents and voltages for the series-feed circuit of Fig. 10-13).[2] A study of these curves will reveal a number of interesting points:

1. The plate current flows only during the time that the plate voltage is at or near its minimum so that the loss on the plate is comparatively low. This results in an efficiency, not including the filament or grid losses, of 70 to 85 per cent. It also makes the plate current dependent *almost entirely on the minimum plate voltage* $e_{b\,min}$, regardless of whether the crest value of the alternating plate voltage E_{pm} is high or low.

2. The plate and grid voltages are essentially sinusoidal, since they are developed across resonant circuits that have a reasonably high Q (where $Q = \omega L/R$). The load circuit presents no reactive component when tuned to resonance, and the grid and plate voltages differ in phase by exactly 180 deg as shown.

3. The maximum positive value of the grid voltage ($e_{c\,max}$) is approximately equal to the minimum value of the plate voltage ($e_{b\,min}$). If the grid were made *more* positive than the plate, the plate-current curve would have a dip at the top, giving it a hat-shaped appearance, and the peak grid current would be considerably increased. This would cause an appreciable increase

[1] For definition of the word *tank*, as used here, see Appendix A, p. 606. Definitions of other unfamiliar terms may also be found in Appendix A.

[2] These curves are presented here without derivation to familiarize the reader with the manner of operation of class C amplifiers. Methods of derivation are presented in a later section.

in grid losses; and the plate losses would also be considerably increased, since the plate-current peaks would occur on either side of the minimum plate voltage where the voltage was somewhat above the minimum. Thus a decrease in efficiency will result if the grid voltage is increased much above the minimum plate voltage. On the other hand, reducing the maximum positive grid voltage much below the minimum plate voltage would cause a pronounced decrease in the average plate current and a falling off of the power input and therefore of the power output.

4. Grid current flows during the time that the grid voltage is positive, so that appreciable power is required to drive the tube as contrasted with the almost zero power required by class A_1 amplifiers.

5. The power flow to the resonant circuit is not continuous. This may be seen by noting that this power is $-e_pi_b$, shown as one of the power curves in Fig. 10-15.[1] Power is delivered to the output circuit only during the interval of plate-current flow, with zero power output from the tube for the remainder of the cycle. On the other hand the power demanded by the load is given by $e_p^2/R_{(\omega_0)}$, where $R_{(\omega_0)}$ is the equivalent resistance presented by the tuned circuit at resonance. The curve of power output determined by this relation is plotted in Fig. 10-16 together with a replot of the power delivered to the load circuit $-e_pi_b$. It is evident that the instantaneous power delivered to the load circuit is not always equal to the power demanded by the load, and the stored energy in the capacitance and inductance of the tuned circuit must supply the difference. The areas crosshatched with a negative slope represent the energy drawn *from* storage; those crosshatched with a positive slope represent the additional energy delivered to the tank circuit over and above the instantaneous load demand and therefore put *into* storage to supply the output during periods of no plate-current flow. The total areas of each

[1] The alternating voltage across the tank circuit e_p is normally considered as acting in a positive direction toward the plate, as shown by the plus and minus signs in Fig. 10-13. To find the power *absorbed* by a circuit element, the voltage across its terminals must be multiplied by the current that *enters its positive terminal*, which in this case is the current $-i_b$. Thus the power delivered to the tank circuit is $e_p(-i_b) = -e_pi_b$. It is obvious that the same conclusion may be reached by taking the product of the current i_b and the voltage of the tank circuit *that opposes this current*. This voltage is evidently $-e_p$, and the product of this current and voltage is again $-e_pi_b$.

of the two types of crosshatching must necessarily be equal. The noncrosshatched areas under the curves evidently represent energy delivered directly from tube to load without going through storage.

This need for storage is one of the prime requisites of class C

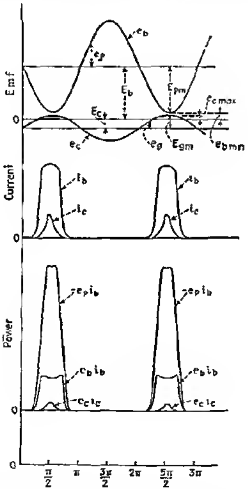


FIG. 10-15. Wave shapes of currents, voltages, and power flow in a class C amplifier with series feed

amplifiers if good wave shape is desired. If the energy stored is large compared with the energy demanded by the load, the relative variations in stored energy throughout the cycle will be small and the wave shape of the plate voltage will be very nearly sinusoidal; but if the energy stored is not much greater than the demand, the stored energy will vary widely throughout the cycle and the wave shape will be poor. This will be further demonstrated in the design of the tank circuit of a class C amplifier presented in a later section.

Wave Shapes in Class C Amplifiers. Parallel Feed. The curves for the parallel-feed circuit of Fig. 10-14 are slightly different and are shown in Fig. 10-17. The plate and grid voltages, currents, and losses are seen to be the same, but the instantaneous

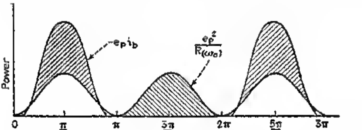


FIG. 10-16. Illustrating the need for energy storage in the tank circuit of a series-feed class C amplifier.

power flow to the tank circuit is somewhat altered. Also the currents i_{rp} and i_{cb} are plotted. The significance of these current symbols is indicated in Figs. 10-14 and 10-18. Figure 10-18a is the equivalent circuit for alternating components only, where R_L represents the effect of the power consumed by the output circuit, and Fig. 10-18b shows direct components only. The use of arrows to indicate the positive direction of current flow conforms with the material in Appendix E. The assumed positive direction of current flow for i_b, i_p, and I_b is in accord with previous work and with the notation of Appendix A. The directions assumed for the other currents are purely arbitrary but logically conform to the idea of an equivalent generator $-\mu E_g$ within the tube.[1]

[1] It may be argued with considerable truth that since the plate current is very far from sinusoidal, it is not possible to use the equivalent generator theorem in the analysis of class C amplifiers. It is possible, however, to obtain much information of value by considering the fundamental com-

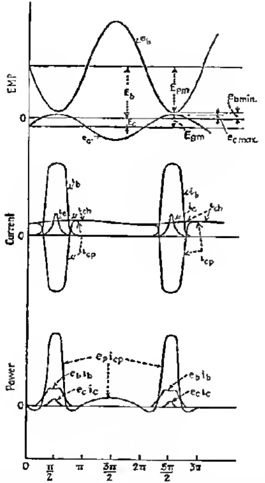


FIG. 10-17. Wave shapes of currents, voltages, and power flow in a parallel-feed class C amplifier.

ponents only of the currents and voltages and then applying the equations developed in earlier chapters. By so doing it is possible, for example, to draw vector diagrams when desired and thus show the effect of phase shift and other factors.

The power delivered to the tank circuit of a parallel-feed amplifier is $e_p i_{cp}$ (the a-c drop through C_p being normally negligible), and the power output is again e_p^2/R_{eq}, where R_{eq} is the equivalent resistance presented by the tank circuit at resonance. These two curves are shown in Fig. 10-19, where the crosshatching has the same significance as in Fig. 10-16. In this type of circuit the flow of power to the tank circuit is much more uniform than in the series-feed circuit, owing to the energy storage in the blocking condenser and choke.

In spite of the difference in the manner in which energy is supplied to the tank circuit in series- and parallel-feed circuits the

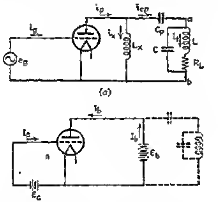


FIG. 10-18. Equivalent circuits of Fig. 10-14; (a) is for alternating components only and (b) for the direct components.

performance of the two is virtually the same; and although the remainder of the discussion in this chapter will be confined almost entirely to the more commonly used parallel-feed circuit, the equations and conclusions will apply almost equally well to those employing series feed.

Power Required to Drive a Class C Amplifier. The instantaneous power P_{dr} supplied to the input circuit by the driving source is equal to the product of the voltage of this source e_g and the current flowing i_c, or

$$P_{dr} = e_g i_c = (e_c - E_c) i_c = e_c i_c - E_c i_c \qquad (10\text{-}28)^*$$

The term $e_c i_c$ is the instantaneous power absorbed by the grid

losses; the term $(-E_c i_c)$ is the instantaneous power absorbed by the grid-bias source. That this term represents power absorbed, not delivered, may be seen by recalling that E_c is always negative, so that the numerical value of the term in the parentheses is positive, denoting the absorption of power. This conclusion may also be reached by noting that the flow of grid current is in such a direction as to charge any battery that might be used for grid-bias purposes. Therefore the driving source for a class C amplifier must have sufficient power capacity to supply both the grid losses and the power absorbed by the bias source.

Design of Class C Amplifiers. Exact Method. Unlike class A amplifiers the d-c power input to class C amplifiers varies with the output, and it is possible to adjust the load resistance and the d-c

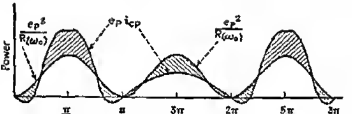


FIG. 10-19. Illustrating the need for energy storage in the tank circuit of a parallel-feed, class C amplifier.

potentials to give maximum efficiency. Since the principal reason for using class C amplifiers is the higher efficiency obtainable, it is desirable to determine the operating conditions for maximum efficiency at any given power output. No simple set of equations can be developed toward such a solution, since the plate current is nonsinusoidal (Fig. 10-12), hence a cut-and-try method must be used. The most thorough method of handling this problem is one in which the solution is obtained directly from the static characteristic curves by step-by-step methods.[1] A number of different combinations of the independent variables are selected, the neces-

[1] For a detailed discussion of such a method, see D. C. Prince, Vacuum Tubes as Power Oscillators, *Proc. IRE*, 11, pp. 275, 405, 527, June, August, October, 1923; also, D. C. Prince and F. B. Vogdes, Vacuum Tubes as Oscillation Generators, *Gen. Elec. Rev.*, 30, pp. 320, 501, June, October, 1927; and 31, p. 97, February 1928.

sary computations performed, and a set of curves plotted from the results. From these curves correct alternating and direct tube voltages may be obtained to give any desired output at maximum efficiency. With this information the tank-circuit design may be carried out in accord with the material beginning on page 414.

An exact solution is seldom used in practice, as it is extremely laborious, and approximate methods give results of sufficient accuracy for commercial purposes.

Design of Class C Amplifiers Using Triode Tubes. Approximate Method.[1] An approximate design of a class C amplifier may be facilitated by assuming the plate-current pulse of Fig. 10-17 to be a portion of a sine wave of the same frequency as that of the driving source. Wagener has shown, by analysis of the actual plate-current pulses of a number of typical amplifiers, that such an assumption does not introduce very appreciable errors. Such an assumption permits computation of the ratio of the peak current to the average (or direct) current and of the fundamental component to the average current for various angular widths of the plate-current pulse, using the general method outlined in Appendix C. The results of such an analysis are shown in the curves of Fig. 10-20, and the significance of the angle θ_p and of the peak plate current $i_{b\ max}$ is indicated in Fig. 10-21. The fundamental component of current I_p in Fig. 10-20*b* is given in rms values.

The grid-current pulse is somewhat different in shape from that of the plate current, and Wagener has shown that it may be assumed to be a sine-squared function. Under this assumption it may be analyzed in the same manner as was the plate current, and these results are also plotted in Fig. 10-20, with the significance of the terminology indicated in Fig. 10-21.

It is now possible to carry out a very satisfactory design with the aid of the curves of Fig. 10-20 and of the static characteristic curves of the tube. The procedure calls for assuming an angle

[1] A number of approximate methods have been developed. The one presented here is essentially that proposed by W. G. Wagener, Simplified Methods for Computing Performance of Transmitting Tubes, *Proc. IRE*, 25, p. 47, January, 1937. For other methods see I. E. Mouromtseff and H. N. Kozanowski, Analysis of the Operation of Vacuum Tubes as Class C Amplifiers, *Proc. IRE*, 23, p. 752, July, 1935; Frederick Emmons Terman and Wilber C. Roake, Calculation and Design of Class C Amplifiers, 24, p. 620, April, 1936; and W. L. Everitt, Optimum Operating Conditions for Class C Amplifiers, *Proc. IRE*, 22, p. 152, February, 1934.

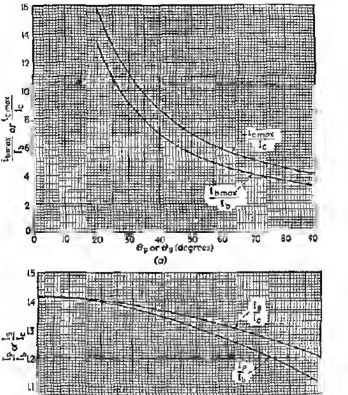


Fig 10-20 Curves based on the Fourier series analysis of theoretical plate- and grid-current wave shapes, as developed by Wagener (I_p and I_g are in rms values)

θ_p and a direct plate voltage and current to give approximately the desired output within the limits of the tube ratings After the design is completed for a given θ_p if the output, efficiency, and driving power are not as desired, a new value of θ_p may be assumed and the design reworked until the results are satisfactory.

Thus the assumption of θ_p involves more of a cut-and-try process than an out-and-out guess. The angle θ_p should be somewhat less than 90 deg if the tube is to be a class C amplifier but not too much less if low driving power is desired. About 75 deg is usually a good value.

It is also possible to repeat the design for other values of direct plate voltage and plate current to secure the best results. However the direct plate voltage is usually made nearly equal to the maximum rating of the tube, as increasing the plate voltage always results in higher efficiency. As to the plate current, there is usually sufficient latitude due to variations in individual tube characteristics and the possibility of minor adjustments in the completed circuit to make unnecessary more than one or two trials. Thus the procedure usually calls for a plate voltage as near to the maximum rating as the available plate supply will permit and a direct plate current that will give the desired output with an efficiency of the order of 75 per cent (a reasonable figure for class C operation) except that the plate current must not be so large as to cause plate dissipation in excess of the tube rating.

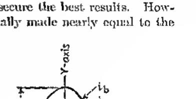


FIG. 10-21. Section of the plate- and grid-current curves of a class C amplifier indicating the significance of the terminology.

Having assumed values of E_b, I_b, and θ_p under the limitations of the preceding paragraphs the steps to be followed are:

1. The peak plate current $i_{b\,max}$ is found by multiplying I_b by the ratio read from the curve of Fig. 10-20a for the assumed θ_p.

2. The minimum plate voltage ($e_{b\,min}$), and the maximum positive grid voltage ($e_{c\,max}$) must be approximately equal in triodes for most efficient operation.[1] The correct voltages are found from the static characteristic curves to give the value of $i_{b\,max}$ obtained in step 1.

3. The alternating plate voltage is found by noting from Fig. 10-15 or Fig. 10-17 that $E_{pm} = E_b - e_{b\,min}$. Multiplying E_{pm} by 0.707 gives the rms plate voltage E_p.

[1] See the discussion of Fig. 10-15 on p. 398. A discussion of these voltages in pentodes is given on p. 413.

4. The rms plate current I_p is found by multiplying I_b by the ratio read from the curve of Fig. 10-20*b* for the assumed θ_p.
5. The power output from the tube is $P_0 = E_p I_p$.
6. The power input to the plate is $P_i = E_b I_b$.
7. The plate efficiency $= P_0/P_i$.
8. The grid bias is computed by means of an equation involving previously determined factors. To derive this equation consider Fig. 10-22 which is a section of the plate- and grid-voltage curves of Fig. 10-15 or 10-17 with the Y axis placed at the point of minimum plate voltage to correspond with Fig. 10-21. The angle θ_g is the angle at which the grid current drops to zero and is, therefore, equal to that value of ωt at which the instantaneous grid voltage is zero, since no appreciable grid current will flow with a negative grid voltage. The angle θ_p is the angle at which the plate current drops to zero and is somewhat larger than θ_g, since the plate current will not go to zero until the grid voltage is equal to its cutoff value, or

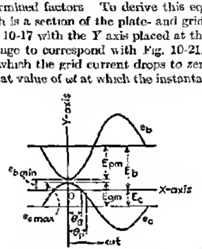


FIG. 10-22. Section of the plate- and grid-voltage curves of a class C amplifier indicating the significance of the terminology.

$$e_{c0} = -\frac{e_{b0}}{\mu} \tag{10-29}$$

where e_{c0} and e_{b0} are the values of e_c and e_b, respectively, at $\omega t = \theta_p$. Since $e_c = E_c + E_{gm}\cos\omega t$ and $e_b = E_b - E_{pm}\cos\omega t$, we may also write[1]

$$e_{c0} = E_c + E_{gm}\cos\theta_p \tag{10-30}$$

$$e_{b0} = E_b - E_{pm}\cos\theta_p \tag{10-31}$$

[1] The minus sign must be used with the plate voltage to conform to Fig. 10-22, the plate voltage being out of phase with the grid voltage by 180 deg.

which are the equations of the instantaneous grid and plate voltages at $\omega t = \theta_p$.

We may also write from Fig. 10-22[1]

$$E_{pm} = E_b - e_{b\,\min} \tag{10-32}$$

$$E_{gm} = e_{c\,\max} - E_c \tag{10-33}$$

Substitution of Eqs. (10-32) and (10-33) in Eqs. (10-30) and (10-31), to eliminate E_{pm} and E_{gm}, and substitution of the resulting values of $e_{b\theta}$ and $e_{c\theta}$ in Eq. (10-29) will give an equation that may be solved for E_c:

$$E_c = -\frac{1}{1-\cos\theta_p}\left[\frac{E_b}{\mu}(1-\cos\theta_p) + \left(\frac{e_{b\,\min}}{\mu} + e_{c\,\max}\right)\cos\theta_p\right] \tag{10-34}$$

All the terms on the right-hand side of this equation are known from preceding steps, so that E_c may be computed from this equation.

9. The alternating grid voltage may now be determined from Eq. (10-33), the rms grid voltage E_g being 0.707 E_{gm}.

10. The grid angle θ_g must be determined. Since the grid voltage is zero at $\omega t = \theta_g$ (see Fig. 10-22), we may write the grid-voltage equation for $\omega t = \theta_g$ as

$$e_{c\,(\text{at}\,\omega t=\theta_g)} = E_c + E_{gm}\cos\theta_g = 0$$

or

$$\cos\theta_g = -\frac{E_c}{E_{gm}} \tag{10-35}$$

The angle θ_g may be found from this equation.

11. The peak grid current $i_{c\,\max}$ is found from the static characteristic curves for the values of $e_{b\,\min}$ and $e_{c\,\max}$ found in step 2.

12. The direct grid current I_c is found by reading the ratio $i_{c\,\max}/I_c$ from Fig. 10-20*a* for the value of θ_g found in step 10 and then solving for I_c.

13. Finally the grid driving power must be determined. It

[1] It must be remembered that in Eq. (10-33) the bias voltage E_c is numerically negative, so that although E_{gm} is actually the numerical sum of the other two voltages as indicated by Fig. 10-22, it must be expressed as the algebraic difference in Eq. (10-33).

may be found from Eq. (10-28)[1] or it may be found by noting that all the a-c power for driving the grid must be supplied by the source of alternating voltage impressed between grid and cathode. From Figs. 10-13 and 10-14 it may be seen that this power is equal to the product of the alternating grid voltage and current or

$$P_{dr} = E_g I_g \qquad (10\text{-}36)$$

The voltage E_g was found in step 9 and I_g may be found from the upper curve of Fig. 10-20*b*.

Example. As an example of the foregoing design procedure assume that an amplifier is to be designed using an 833 tube. The maximum ratings of this tube, when used as a class C amplifier without modulation, are given by the manufacturer as

$$E_b = 3000 \text{ volts}, E_c = -500 \text{ volts}, I_b = 500 \text{ ma}, I_c = 75 \text{ ma}$$

$$\text{Plate dissipation} = 300 \text{ watts}$$

Assume that the tube is to be operated at 3000 volts with the maximum output obtainable without exceeding the plate dissipation. At an efficiency of 75 per cent the input power for 300 watts dissipation would be 1200 watts, requiring a plate current of 0.4 amp. Let this be the assumed plate current, and let the angle θ_p be 75 deg.

Step 1. From the curve of Fig. 10-20*a*, $i_{b\,max}/I_b = 3.8$. Therefore $i_{b\,max} = 3.8 \times 0.4 = 1.52$ amp.

Step 2. The static characteristic curves of an 833 tube are given in Figs. 10-23 and 10-24. On the plate-current characteristics of Fig. 10-23 a curve has been drawn which is the locus of currents for equal grid and plate voltages ($e_b = e_c$). Since $e_{b\,min} = e_{c\,max}$, these two voltages may be read from this locus curve for a current of 1.52 amp and are found to be 160 volts.

Step 3. $E_p = 0.707(3000 - 160) = 0.707 \times 2840 = 2000$ volts.

Step 4. From the curve of Fig. 10-20*b* $I_p/I_b = 1.19$. Therefore $I_p = 1.19 \times 0.4 = 0.476$ amp.

Step 5. $P_o = 2000 \times 0.476 = 952$ watts.

Step 6. $P_s = 3000 \times 0.4 = 1200$ watts.

Step 7. $\text{Plate efficiency} = \dfrac{952}{1200} \times 100 = 79.6$ per cent.

Step 8. $$E_c = -\frac{1}{1 - 0.26}\left[\frac{3000}{35}(1 - 0.26) + \left(\frac{160}{35} + 160\right)0.26\right]$$

$$= -144 \text{ volts}$$

−143.5366

[1] For one method of solution, see H. P. Thomas, Determination of Grid Driving Power in Radio-frequency Amplifiers, *Proc. IRE*, 21, p. 1134, August, 1933.

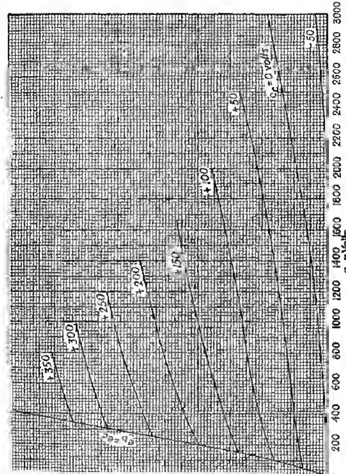


FIG. 10-23. Static characteristic curves of an 833 tube; plate current.

Step 9 $E_s = 0.707(160 + 141) = 0.707 \times 301 = 215$ volts

Step 10 $\cos\theta_s = -\frac{-141}{301} = -0.471,\ \theta_s = 61.5$ deg.

Step 11 From the grid-current static characteristic curves of Fig. 10·21 the peak grid current $i_{c\,max}$ is found for $e_{b\,min} = 160$ and $e_{c\,max} = 160$. It is 0.41 amp

Step 12 $I_c = 0.41/5.7 = 0.077$ amp, where 5.7 is obtained from Fig 10·20*a*

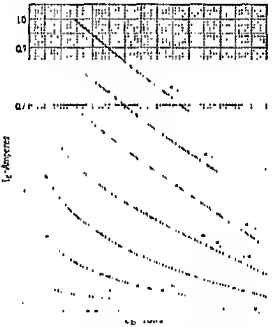


FIG 10-21 Static characteristic curves of an 833 tube, grid current.

Step 13. $I_g = 1.31 \times 0.077 = 0.101$ amp, where 1.31 is obtained from Fig. 10·20*b*. Therefore, $P_{dr} = 215 \times 0.101 = 21.7$ watts

The plate dissipation is the difference between the output of 932 watts and the input of 1200 watts and is less than the maximum permissible dissipation for this tube. It would, therefore, be possible to redesign the amplifier using a slightly higher direct plate current and so increase its output

The direct grid current from step 12 is a little greater than the maximum rating. Theoretically this should require a slight decrease in the driving voltage E_g which will in turn decrease the power output. Actually the direct current and driving power of a tube vary rather widely in practice owing in part to the wide range of impedance found in practical driving circuits. In general, computations of grid current and driving power by this method give results that are too high, being fortunately the desirable side on which to err.

It may be of interest to repeat this example for another value of θ_p. Assume $\theta_p = 65$ deg using the same values of E_b and I_b. The results of each step are given below, but the actual computations are omitted for brevity.

Step 1. $i_{b\,\max} = 1.7$ amp.
Step 2. $e_{b\,\min} = e_{c\,\max} = 180$ volts.
Step 3. $E_p = 1990$ volts.
Step 4. $I_p = 0.497$ amp.
Step 5. $P_o = 990$ watts.
Step 6. $P_i = 1200$ watts.
Step 7. Plate efficiency $= 82.5$ per cent.
Step 8. $E_c = -222$ volts.
Step 9. $E_g = 285$ volts.
Step 10. $\theta_g = 56.5$ deg.
Step 11. $i_{c\,\max} = 0.48$ amp.
Step 12. $I_c = 0.070$ amp.
Step 13. $P_{dr} = 20.9$ watts.

A comparison of these results with those for $\theta_p = 75$ deg shows the alternating plate voltage to be virtually unchanged. This is characteristic of class C amplifiers because the minimum plate voltage cannot vary widely without causing a very large change in plate current. (The plate current flows only at or near the time of minimum plate voltage, and its amplitude is therefore almost directly proportional to $e_{b\,\min}$.) Also the output is seen to have increased slightly with an appreciably higher efficiency. The efficiency is always increased by a decrease in the angle θ_p, since this decreases the average plate voltage impressed during the conducting period. At the same time the peak plate current $i_{b\,\max}$ and the driving power P_{dr} are increased. The increase in driving power is particularly bad and may make the use of a larger angle θ_p preferable, even at a sacrifice in efficiency to keep the driving power within desired limits.

Design of Class C Amplifiers Using Pentodes and Tetrodes.[1] The foregoing approximate design procedure may be applied to pentodes and tetrodes with a few minor modifications and will give reasonably accurate results. In step 2, $e_{b\,\min}$ in pentodes is made only slightly more positive than the suppressor grid poten-

[1] See also F. E. Terman and J. H. Ferns, The Calculation of Class C Amplifier and Harmonic Generator Performance of Screen Grid and Similar Tubes, *Proc. IRE*, **22**, p. 359, March, 1934.

tial (which is commonly slightly positive with respect to the cathode in pentodes used for class C service). More particularly $e_{b\,min}$ should be equal to the lowest voltage at which the plate current is still nearly independent of changes in plate voltage. For the pentode to which the curves of Fig. 9-49 (page 334) apply, $e_{b\,min}$ should be about 60 volts (the suppressor voltage for this tube was zero). The grid voltage $e_{c\,max}$ should then be made sufficiently high to cause the desired $i_{b\,max}$ (found from step 1) to flow.

In step 8 the cutoff voltage e_{c0} in a pentode is equal to $-E_{c2}/\mu_{sgp}$, where E_{c2} is the screen-grid voltage and μ_{sgp} is the mu factor of the screen grid relative to the control grid with the plate current being held constant. Substituting this relationship and that of Eq. (10-33) into Eq. (10-30) gives for the grid bias

$$E_c = \frac{-(E_{c2}/\mu_{sgp}) - e_{c\,max}\cos\theta_p}{1 - \cos\theta_p} \qquad (10\text{-}37)$$

The mu factor μ_{sgp} is not usually given in tube-data sheets, but the cutoff voltage $-E_{c2}/\mu_{sgp}$ may be estimated, with some degree of accuracy, from the static characteristic curves by assuming the dynamic curve to be reasonably straight. Thus in Fig. 9-49 the cutoff voltage would be about -35 volts if the curves for the various grid voltages were equally spaced; *i.e.*, if the dynamic curve were straight.

Similar relations hold for a tetrode except that the minimum plate voltage should be slightly more positive than the screen grid, rather than the suppressor.

Design of the Tank Circuit to Produce Correct Plate Voltage. The alternating plate voltage computed by the method of the preceding sections must be secured in an actual amplifier by proper design of the tank circuit, *i.e.*, by choosing suitable values of L, C, and R_L (Fig. 10-18). This design may be carried out with the aid of the equivalent a-c circuit of Fig. 10-18a, for which

L = inductance of tank circuit
C = capacitance of tank circuit
R_L = equivalent series resistance in tank circuit representing both power output and coil losses

We may now write

$$P_0 = I_t^2 R_L \qquad (10\text{-}38)$$

where I_t is the rms value of the current flowing through the tank

from the amplifier specifications or may be computed by the design procedure of the preceding sections, except Q. The term Q essentially determines the wave shape and stability of the amplifier and must be chosen to fit the conditions under which the amplifier is to operate. It is usually chosen by experience, being about 12 to 15 for highest efficiency with reasonably good wave shape, and somewhat larger if very good wave shape is desired at some sacrifice in efficiency An analysis of its effect on wave shape is given in a later section, page 419.

Energy Storage in Tank Circuits. The final design of the tank circuit will also depend upon the amount of harmonic content that may be permitted, which is in turn a function of the total energy stored in the tank inductance and capacitance. Figures 10-16 and 10-19 showed that the supply of power to the tank circuit by the tube was periodic and that some of the output must be supplied from storage during certain portions of the cycle. The effect of this storage may be likened to the effect of energy storage in the flywheel of a reciprocating engine. In such an engine the constancy of the angular velocity of the drive shaft, and therefore the purity of the sine-wave motion of the crosshead, is a direct function of the size of the flywheel. Similarly, the amount of energy stored in the tank circuit of a class C amplifier determines the purity of the output wave. In particular, as illustrated more fully in the two ensuing sections, the energy storage must exceed, by an appreciable amount, the energy dissipated during a cycle if the wave shape is to be satisfactory for most purposes

The total energy stored in the tank circuit is the sum of the storage in the inductance and in the capacitance, or

$$\text{Stored energy} = \frac{Li_t^2}{2} + \frac{Ce_t^2}{2} \qquad (10\text{-}45)$$

where i_t = tank current flowing through the inductance

e_t = voltage across tank circuit between points a and b, Fig. 10-18, being equal to e_p for most practical purposes

If the energy storage is sufficient to keep the harmonics small, the voltage across the condenser is essentially sinusoidal, or

$$e_t = E_{tm} \sin \omega_0 t \qquad (10\text{-}46)$$

The current flowing through the tank inductance is found by

dividing this voltage by the series impedance of L and R_L (Fig. 10-18a). In a properly designed tank circuit $R_L \ll \omega_0 L$, and we may write

$$i_t = -\frac{E_{tm}}{\omega_0 L} \cos \omega_0 t \qquad (10\text{-}47)$$

where the minus sign is used because the current in an inductive circuit lags the impressed voltage.

Substituting Eqs. (10-46) and (10-47) into Eq. (10-45) gives

$$\text{Stored energy} = \frac{L E_{tm}^2}{2(\omega_0 L)^2} \cos^2 \omega_0 t + \frac{C E_{tm}^2}{2} \sin^2 \omega_0 t \qquad (10\text{-}48)$$

At resonance $\omega_0 L = 1/\omega_0 C$, and substituting the value of L, obtained from this relation, into the preceding equation gives

$$\begin{aligned} \text{Stored energy} &= \frac{C E_{tm}^2}{2} \cos^2 \omega_0 t + \frac{C E_{tm}^2}{2} \sin^2 \omega_0 t \\ &= \frac{C E_{tm}^2}{2} = C E_t^2 \end{aligned} \qquad (10\text{-}49)^*$$

from which it is apparent that the stored energy is increased by using a larger condenser in the tank circuit and, therefore, to maintain the same resonant frequency, a smaller inductance.

The energy storage should be kept as low as possible consistent with securing the desired purity of output wave, since an increase in tank current and, therefore, an increase in tank coil losses always accompanies any increase in energy storage. Increased coil loss necessarily lowers the over-all efficiency of the amplifier and should be avoided as far as possible.

Design of the Tank Circuit to Secure Sufficient Energy Storage. The problem of designing the tank circuit to secure sufficient energy storage may be handled in a number of ways. A logical procedure is to supply storage sufficient to keep the second harmonic voltage within a certain percentage of the fundamental. To carry out this method of attack it is first necessary to determine the impedance of the tank circuit at both frequencies.

The impedance of the tank circuit at any frequency $\omega/2\pi$ may be determined by writing the impedance $\mathbf{Z}$ between points a and b of Fig. 10-18a as the reciprocal of the sum of the reciprocals of the impedance of the two branches of the tank circuit

$$\mathbf{Z} = \frac{1}{j\omega C + \dfrac{1}{R_L + j\omega L}} = \frac{R_L + j\omega L}{1 + j\omega R_L C - \omega^2 LC} \tag{10-50}$$

If $\omega = 2\omega_0$ (where ω_0 is 2π times the frequency of resonance) Eq (10-50) gives for the impedance to the second harmonic

$$\mathbf{Z} = \mathbf{Z}_{(2\omega_0)} = \frac{R_L + j2\omega_0 L}{1 + j2\omega_0 R_L C - 4\omega_0^2 LC} \tag{10-51}$$

But at resonance $\omega_0 L = 1/\omega_0 C$, or

$$C = \frac{1}{\omega_0^2 L} \tag{10-52}$$

Substituting this value of C into Eq. (10-51) gives

$$\mathbf{Z}_{(2\omega_0)} = \frac{R_L + j2\omega_0 L}{1 + j\dfrac{2R_L}{\omega_0 L} - 4} \tag{10-53}$$

or

$$\mathbf{Z}_{(2\omega_0)} = \omega_0 L \frac{R_L + j2\omega_0 L}{-3\omega_0 L + j2R_L} \tag{10-54}$$

The *magnitude* of $\mathbf{Z}_{(2\omega_0)}$ is

$$Z_{(2\omega_0)} = \omega_0 L \sqrt{\frac{R_L^2 + 4\omega_0^2 L^2}{9\omega_0^2 L^2 + 4R_L^2}}$$

and since $\omega_0 L \geqq 12R_L$ in a properly designed amplifier, the R_L^2 terms may be neglected Equation (10-54) may then be rewritten by neglecting the R_L terms

$$\mathbf{Z}_{(2\omega_0)} = -j\frac{2\omega_0 L}{3} \tag{10-55)*}$$

In a similar manner the impedance at ω_0 may be found by replacing ω in Eq. (10-50) with ω_0 and simplifying by the further substitution of Eq. (10-52) and by neglecting the R_L^2 terms. This gives

$$\mathbf{Z}_{(\omega_0)} = \frac{(\omega_0 L)^2}{R_L} \tag{10-56)*}$$

The absence of a reactive term in Eq. (10-56) shows that the tank impedance is purely resistive at the fundamental frequency, a natural result of operating the circuit at resonance.

The voltage set up across the tank circuit at the frequency of the fundamental is equal to the product of the fundamental component of the plate current times the impedance of the tank circuit to the fundamental, or, in magnitude only

$$E_{t(\omega_0)} = I_p Z_{(\omega_0)} = \frac{(\omega_0 L)^2}{R_L} I_p \tag{10-57}$$

and the voltage set up at the second harmonic of the impressed frequency is, in magnitude only

$$E_{t(2\omega_0)} = aI_p Z_{(2\omega_0)} = \frac{2a\omega_0 L}{3} I_p \tag{10-58}$$

where a is the ratio of the magnitudes of the second-harmonic and fundamental components of the plate current, and I_p is the fundamental component. Let k be the ratio of the magnitudes of the second-harmonic and fundamental components of the voltage across the output circuit; *i.e.*, $k = E_{t(2\omega_0)}/E_{t(\omega_0)}$. Then

$$k = \frac{2a\omega_0 L I_p/3}{(\omega_0 L)^2 I_p/R_L} = \frac{2a}{3}\frac{R_L}{\omega_0 L} \tag{10-59}$$

This equation may be solved for $\omega_0 L/R_L = Q$ to give

$$Q = \frac{2a}{3k} \tag{10-60*}$$

The value of L required to give the correct plate voltage with the given harmonic content may now be determined by substituting Eq. (10-60) into (10-43), since all terms in the latter equation were known except Q and L. C may then be determined from Eq. (10-44).

The blocking condenser C_p (Fig. 10-14) need be only of sufficient capacity to cause negligible drop in voltage between the plate of the tube and the tank circuit. Similarly, the reactance of the choke L_x need be only large as compared with the impedance of the tank circuit $Z_{(\omega_0)}$ at the fundamental frequency. Consequently, no special design procedure is required for their evaluation.

Significance of Q in the Design of Class C Amplifiers. The factor a in Eq. (10-60) is a function of the wave shape of the plate current and is nearly the same for all class C amplifiers no matter what may be the percentage of second harmonic in the output circuit. Thus Q is a direct measure of the harmonic content in

the output voltage, and it is common practice to specify the minimum Q that will maintain the harmonic content within the desired limits. Most amplifiers are, therefore, designed not by determining Q from Eq. (10-60) but by specifying Q from experience records. A minimum value for Q may be said to be about 12.5, which corresponds roughly to a harmonic content of 3 per cent, but Q may be made much larger where purity of wave is essential.[1]

Another concept of the factor Q is formed by taking the ratio of the energy stored per cycle to the energy dissipated per cycle. The energy stored was given by Eq (10-49), while the energy dissipated per cycle is equal to $I_t^2R_L$ multiplied by the time of 1 cycle, $1/f_0$.

$$\frac{\text{Energy stored per cycle}}{\text{Energy dissipated per cycle}} = \frac{CE_t^2}{I_t^2R_L/f_0} \tag{10-61}$$

We may also write $E_t = \omega_0 LI_t$ (neglecting the small effect of R_L) and $C = 1/\omega_0^2L$ from Eq. (10-52). Therefore Eq. (10-61) may be reduced to

$$\frac{\text{Energy stored per cycle}}{\text{Energy dissipated per cycle}} = \frac{\omega_0 L}{2\pi R_L} = \frac{Q}{2\pi} \tag{10-62*}$$

This equation shows Q to be *a measure of the ratio of the energy stored to the energy dissipated per cycle*, a concept of considerable usefulness.

Q may also be expressed as the ratio of the circuit volt-amperes to the circuit watts. We may write

$$Q = \frac{\omega L}{R_L} = \frac{\omega LI_t^2}{R_LI_t^2} = \frac{EI}{W} \tag{10-63*}$$

where EI = volt-amperes in circuit
W = watts output

Example. Suppose that the tank circuit of a class C amplifier is to be designed from the solution for the 833 tube (page 410) using $\theta_p = 75$ deg. It will be assumed that $Q = 12.5$ will give satisfactory wave shape Sub-

[1] Values of Q of *less* than 12.5 may be necessary in certain cases to avoid phase shift of the side bands at the higher audio frequencies, as when feedback is used. In such cases harmonic filters must be inserted in the output to purify the wave.

stituting the values of P_0 and E_p computed on page 410 into Eq. (10-42) gives

$$\omega_0 L = \frac{(2000)^2}{952 \times 12.5} = 336 \text{ ohms}$$

which, from Eq. (10-52), is also $1/\omega_0 C$.

The impedance of the blocking condenser should be small as compared with the impedance of the tank circuit at resonance, but the reactance of the choke should be large as compared with this impedance. From Eq. (10-56) the tank circuit impedance is $(\omega_0 L)^2/R_L = \omega_0 LQ$, or

$$Z_{(\omega_0)} = 336 \times 12.5 = 4200 \text{ ohms}$$

The reactance of the blocking condenser should not be more than about 10 per cent of this resistance, or about 400 ohms, and the reactance of the choke should not be less than about ten times $Z_{(\omega_0)}$ or about 40,000 ohms.

If the frequency at which the amplifier is to operate is 1000 kc, the capacitances and inductances of the circuit may be solved for

$$L = \frac{336}{2\pi 10^6} \times 10^6 = 53.5\,\mu\text{h}$$

$$C = \frac{1}{2\pi 10^6 \times 336} \times 10^{12} = 473\,\mu\mu\text{f}$$

$$C_p \geqq \frac{1}{2\pi 10^6 \times 400} \times 10^{12} \leqq 400\,\mu\mu\text{f}$$

$$L_s \geqq \frac{40{,}000}{2\pi 10^6} \times 10^3 \geqq 6.4\,\text{mh}$$

Approximate Solution for the Tank Circuit of Class C Amplifiers. Many times only a very approximate solution of the class C amplifier is desired, as in the design of laboratory amplifiers or where the power output is low. In such cases final adjustments can be made experimentally after the amplifier has been built, and only a very approximate determination of the tank coil and condenser is required before building the amplifier.

As in the design procedure outlined in the section beginning on page 405 we shall assume that the direct plate voltage and power output are known. Starting with the direct plate voltage we can make a fair estimate of the alternating plate voltage by noting from Figs. 10-15 and 10-17 that the minimum plate voltage is quite small and that the crest value of the alternating plate volt-

age is, therefore, nearly equal to the direct plate voltage. Since the minimum plate voltage is given by

$$e_{b\,\min} = E_b - E_{pm}$$

it is numerically equal to the difference between two large numbers, and a small change in E_{pm} will cause a large change in $e_{b\,\min}$. As stated in item 1, page 398, the plate current is nearly proportional to $e_{b\,\min}$ so that a small change in E_{pm} will cause a large change in plate current and therefore in power input. But the power input cannot vary widely for a given power output, and we must conclude that E_{pm} is nearly constant regardless of the load. We may, therefore, assume that the crest value of the alternating voltage is about 90 to 95 per cent of the direct plate voltage under all conditions of normal operation (In the examples on pages 410 and 413 it was slightly less than 95 per cent.) The rms value may be computed from the crest value and inserted into Eq (10-43) along with the power output and estimated value of Q The solution obtained in this manner will be satisfactory for nearly all cases except those in which L must be predetermined to a high degree of accuracy, as in high-power amplifiers where the investment is large.

Sources of Power for Class C Amplifiers. Plate power for class C amplifiers may be supplied by a battery, a generator, a rectifier operating from the 60-cycle mains, or even directly from a 60-cycle source without rectification. In the last case the tube will produce an output only during the periods of positive supply voltage. The amplifier output will therefore vary throughout the cycle of the a-c supply, being zero for the negative half cycle, increasing to a maximum during the first half of the positive half cycle, and then decreasing again to zero An output wave of this form tunes very broadly and is otherwise unsatisfactory; an unrectified a-c plate supply is therefore used only in emergency or for special applications.

Rectifiers are most commonly used to supply plate potential and may be built to satisfy a wide range of power and voltage requirements Their design was taken up in detail in Chap. 7.

Grid bias is normally obtained from a constant voltage source—as a generator, battery, or rectifier—(known as *fixed bias*); from a cathode resistor as in Fig. 9-39 (page 321) (known as *cathode bias*); or by means of a grid leak (known as *grid-leak bias*) Per-

haps the simplest of these is the grid-leak method in which a suitable resistance is connected in series with the grid, as R_g in Fig. 10-25, thus causing the rectified (or direct) component of grid current to produce a direct voltage for bias purposes. The resistance should be by-passed by a condenser of such size as to offer low impedance to the a-c excitation and thus maintain constant voltage across R_g throughout the cycle of grid excitation voltage. (The actual grid-current flow is of course a series of pulses, as shown in Figs. 10-15 and 10-17.) Such a resistor produces a bias that is a function of the excitation voltage, and any change in excitation is therefore accompanied by an automatic change in bias which maintains approximately correct operating conditions. On the other hand if the excitation should fail, the grid current would drop to zero and the tube would lose its bias

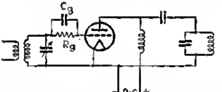


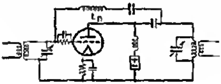
FIG. 10-25. Class C amplifier with grid-leak bias.

entirely. Unless the μ of the tube is very high, such loss of bias is likely to cause the plate current to become so excessive as to seriously damage if not destroy the anode or cathode. Thus protective devices must be used to open the plate circuit in event of excitation failure when grid-leak bias is employed.

Cathode bias is also somewhat self-regulating, since any increase in excitation will produce an increase in average plate current and, therefore, in the bias voltage. Grid-leak bias is sometimes combined with cathode bias, sufficient resistance being inserted in the cathode circuit to prevent damage to the tube in case of excitation failure, with the grid leak supplying the remainder of the bias.

Neutralization. All class C amplifiers using triode tubes are subject to self-oscillation unless proper steps are taken. Inspection of Eq. (9-7) for the input impedance of a triode will disclose that the resistive component becomes negative with an inductive

reactive load (ϕ positive). This simply means that the flow of power will be in the opposite direction; *i.e.*, the grid circuit is capable of supplying useful power output instead of requiring power to drive it. The explanation is that sufficient energy may be supplied from the plate circuit through the grid-plate capacitance to supply all losses in the grid circuit; thus the tube is self-exciting and will oscillate at the frequency to which the grid and the plate circuits are tuned. The inductive reactance in the plate circuit of a class C amplifier needed to produce this effect is provided by a very slight detuning of the tank circuit. Since it is



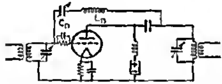


FIG. 10-26 Two simple neutralizing circuits.

impossible to tune the circuit with an accuracy that will ensure resistive or capacitive reaction, it is necessary to prevent oscillation in some other manner.

Perhaps the simplest method of neutralization is to bridge an inductance L_n between grid and plate as in Fig 10-26*a*. The reactance of this coil should equal the reactance of the plate-grid capacitance at the frequency to which the amplifier is tuned, thus setting up a high-impedance parallel resonant circuit between plate and grid and reducing the feed-back current to a negligible quantity. The condenser in series with L_n prevents short-circuiting of the bias and is of low reactance.

The effective reactance of the coil L_n may be varied by inserting a variable condenser in parallel, or in series as in Fig. 10-26*b*. A

series capacitance *must* be used with grid-modulated, class C amplifiers (page 521) to prevent short-circuiting the l-f modulating voltage through the coil L_n. It is of course necessary that the reactance of coil L_n exceed that of condenser C_n when they are in series and be less than that of C_n when in parallel, to produce a net reactance that is inductive in nature.

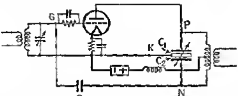


Fig. 10·27. Neutrodyne circuit for neutralizing a tuned amplifier.

Bridge-type Neutralizing Circuits. Another common method of neutralization is to divert the feed-back current by means of a bridge circuit, as in Hazeltine's neutrodyne circuit (Fig. 10-27). The primary of the output transformer is mid-tapped, and the direct plate current is fed into this tap through an r-f choke.[1] The tank condenser is of the split-stator variety with the movable plates (represented by the central rectangular block) grounded. A small neutralizing condenser is inserted between the lower terminal of the tank inductance and the grid of the tube. Proper adjustment of this neutralizing condenser C_n will effectively prevent oscillations.

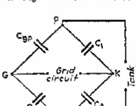


Fig. 10·28. Equivalent bridge circuit of Fig. 10·27.

The operation of this circuit can best be understood with the aid of the equivalent circuit of Fig. 10-28. C_1 and C_2 represent the upper and lower parts, respectively, of the tank condenser; C_{gp} represents the interelectrode capacitance between

[1] The choke is used to avoid the possibility of two resonant frequencies existing in case the coil is not tapped in the exact center or the two halves of the condenser are not identical. It is possible to connect the d-c lead to either end of the tank coil instead of at the center; but this would require a larger choke, since the impressed r-f voltage would be greater.

grid and plate of the tube, and C_n the neutralizing condenser. Any alternating voltages impressed between the grid and cathode are effectively impressed between points G and K and the total tank circuit voltage is impressed between points P and N. The similarity of this circuit to a Wheatstone bridge is at once apparent, where

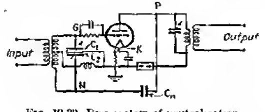


FIG. 10-29 Rice system of neutralization

the tank voltage serves as the source of emf with which the bridge is energized and the grid circuit corresponds to the phones Proper adjustment of C_n will bring the bridge into balance, and zero voltage will be impressed across the grid circuit regardless of the magnitude of the tank voltage.

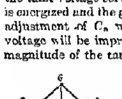


FIG 10-30 Equivalent bridge circuit of Fig. 10-29.

Figure 10-29 shows the Rice system of neutralization It differs from the neutrodyne in that the extra winding for energizing the neutralizing condenser is provided on the grid coil instead of on the plate coil The equivalent circuit (Fig 10-30) is similar to that of Fig. 10-28 except that the grid and plate circuits are interchanged. Since a Wheatstone bridge will balance equally well when the positions of the energizing source and the phones are interchanged, the condenser C_n in this circuit can be adjusted for balance in the same manner as for the circuit of Fig. 10-27.

Where push-pull amplifiers are used, as in Fig. 10-31, the problem of neutralization becomes relatively simple A small condenser is connected between the grid of each tube and the plate of the other tube. In this case the method used is really a combination of both the methods just outlined. C_1, for example, may be considered as neutralizing the upper tube by the neutrodyne method

or as neutralizing the lower tube by the Rice method. Neutralization of a push-pull amplifier by this means is generally more easily accomplished than is neutralization of a single tube so that in very h-f circuits, tubes are nearly always operated in push pull when the power output is of the order of a few watts or higher.

All r-f amplifiers used in radio receivers now employ pentode tubes and therefore do not require neutralization. Many r-f amplifiers used in radio transmitters and in other applications involving appreciable amounts of power also employ pentode (or beam) tubes, but neutralization is frequently desirable anyway, to balance out the effect of stray capacitances in the external circuit as well as any yet remaining within the tube. Furthermore triodes are still used to some extent in high-power amplifiers and,

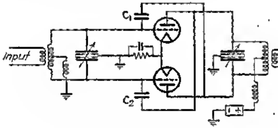


FIG. 10-31. Neutralized push-pull amplifier.

when so used, must be neutralized. Consequently neutralizing circuits continue to be of interest despite the development of screen-grid tubes.

Adjusting Class C Amplifiers. Care must be exercised in adjusting the tuning of class C amplifiers when full plate voltage is applied, especially if grid-leak bias is used. A tank circuit that is off resonance will not set up appreciable alternating voltage, and the instantaneous plate voltage will be excessively high during the period of plate-current flow. This may be seen from Fig. 10-17, which shows that the low instantaneous plate voltage ($e_{b\,min}$) normally existing during the period of current flow is equal to the difference between the direct voltage and the crest value of the alternating voltage; and from Fig. 10-32, which illustrates the high value of $e_{b\,min}$ existing when the tank circuit is detuned and the alternating plate voltage is, therefore, low. Con-

sequently if the tank circuit is off resonance, both the plate voltage and the plate current will be high during the conduction period, and the plate loss will be excessive.

The safe method of tuning a class C amplifier is to make the adjustments with reduced direct voltage on the plate. The tank circuit may then be tuned by adjusting the tank condensers for minimum deflection of the direct-plate-current meter. This method of indicating resonance is possible because the lowest value of $e_{b\,min}$, and therefore of both the instantaneous and the average plate current, occurs at resonance.

The driving circuit may be tuned to resonance by adjusting for maximum reading on the direct-grid-current meter. Tuning of the grid circuit may disturb the performance of the driving amplifier, however, and the adjustment of that amplifier should, therefore, be rechecked.

FIG. 10-32. Illustrating the high values of $e_{b\,min}$ present when the tank circuit of a class C amplifier is detuned.

The proper setting of the neutralizing condenser may be found by removing the plate voltage from the tube and tuning the output circuit to resonance, using a simple type of indicating circuit, such as a thermogalvanometer or even a small lamp or neon tube coupled to the tank inductance. The neutralizing condenser is then adjusted to give minimum response in the indicating device. It is usually necessary to retune the tank circuit and perhaps the input circuit after neutralizing, following which the neutralization condenser must be readjusted. This process should be repeated until no response can be detected in the tank circuit. If complete neutralization cannot be obtained, it probably indicates the presence of external coupling, as inductive coupling between the coils in the plate and grid circuits. This should be eliminated by shielding or by relocation of the parts if the most satisfactory performance of the amplifier is to be realized.

Another method of adjusting the neutralizing condenser, with

the direct plate voltage removed, is to find the position for which the change in the d-c component of the grid current is a minimum as the plate tank tuning condenser is varied through resonance. This change may be indicated by a d-c ammeter in the grid circuit.

Frequency Doublers. It is possible to use a class C amplifier as a frequency doubler by tuning the output circuit to twice the frequency of the driving source. The alternating plate voltage is again sinusoidal but alternates at double frequency, and the plate current is essentially the same as in the conventional class C amplifier, as shown in Fig. 10-33. It may be seen from this figure that the plate-current pulse should be only *half* as wide as that of the normal class C amplifier (Fig. 10-17) if the plate loss per pulse is not to be increased. If the time of current flow is not so reduced, the plate voltage at the beginning and end of the plate-current pulse will be very high, perhaps even in excess of the direct voltage, and very marked increase in plate loss and decrease in efficiency will be the result. Good design therefore calls for an angle of flow $2\theta_p$ nearly proportional to the ratio of input to output frequencies. This means that for the same peak plate current the average current and therefore the power input and power output will be reduced. Thus the maximum output of a frequency doubler is about 60 to 70 per cent of the output of a conventional class C amplifier using the same tube.

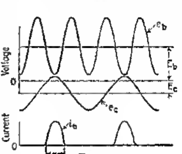


FIG. 10-33. Plate and grid voltages and plate current in a frequency doubler.

Frequency doublers do not require neutralization even when triode tubes are used. The output and input circuits are tuned to different frequencies, and the energy fed back through the grid-plate capacitance will not sustain oscillations.

It is quite possible to tune the output of the amplifier to three, four, or even higher multiples of the driving frequency. However, it is evident from the foregoing discussion that the output at these higher frequencies will be increasingly lower, and it may be more economical to use two frequency doublers, for example,

than one frequency quadrupler. Furthermore, the grid driving losses tend to increase with the degree of multiplication. This is because an increasingly higher bias is required to decrease the width of the plate-current pulse as the ratio of output to input frequency is increased. Thus a tripler should have a much higher bias than a doubler, and reference to Eq. (10-28) will show that this is accompanied by a marked increase in driving power. The only alternative is to reduce the bias, let the current flow throughout a larger angle $2\theta_p$, and operate the tube at a lower plate efficiency. In many cases this must be accompanied by a reduction in amplifier output to prevent excessive plate dissipation.[1]

Frequency doublers are commonly used in connection with crystal oscillators (see page 467) to obtain frequencies higher than those for which crystals may be ground or in constructing radio transmitters that are to operate at two or more frequencies that are multiples of one another.

Class B Linear Amplifiers. *The study of class C amplifiers in the preceding sections of this chapter was based on the assumption of a constant amplitude exciting voltage applied to the grid of the tube.* Many times it is necessary to amplify modulated waves in which the amplitude varies periodically according to a signal the frequency of which is very much less than that of the wave that it modulates.[2] Comparison of Fig. 10-11 for a class B amplifier with Fig. 10-12 for a class C amplifier will show why class B rather than class C amplifiers must be used to amplify modulated waves. If the alternating voltage impressed on the grid of a class C amplifier is the modulated wave of Fig. 10-34a, the amplitude of the plate-current pulses will be zero for an appreciable period on either side of t_1 (Fig. 10-34a). This will produce a plate voltage as in Fig. 10-34c, whereas distortionless operation is represented by the curve of Fig. 10-34b. A class B amplifier, on the other hand, permits the flow of plate current for all values of the impressed alternating voltage, producing a plate voltage as in Fig. 10-34b.

[1] Frequency doublers may be designed by suitably modifying Wagener's method. Other methods are also available such as that given by Robert H. Brown, Harmonic Amplifier Design, *Proc. IRE*, 35, pp. 771–777, August, 1947.

[2] Figure 10-34a represents a modulated wave, a detailed discussion of which will be found in Chap. 13.

The essential difference between class B and class C amplifiers lies in the amount of bias applied, the class B bias being approximately equal to the cutoff voltage, whereas the class C bias may be as much as several times cutoff.[1] Thus the same circuits may be used for class B amplifiers as were described in earlier sections of this chapter for class C amplifiers, and class B design procedure follows the same general lines as class C although somewhat simplified, since E_r is found from the tube characteristics instead of being computed.

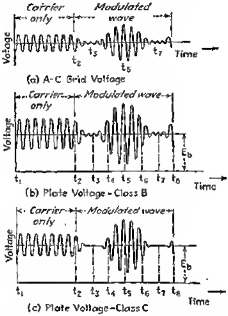


FIG. 10-34. Effect of class B and class C amplifiers on a modulated wave.

There is, however, one very marked difference in the design of class B and C amplifiers. In the latter type the alternating plate voltage is constant, and the amplifier may be designed for maxi-

[1] As may be seen from Fig. 10-11, the bias of a class B amplifier should be a little less than cutoff due to the curvature of the tube characteristic. The correct bias may be found by projecting the essentially straight section of the curve until it intersects the X axis.

mum efficiency, in the former the alternating plate voltage varies from a maximum to zero as the impressed wave varies from a maximum to a minimum over the modulation cycle (Fig. 10-34). Thus a class B amplifier may be designed for maximum efficiency only at the crest of the modulation cycle, and its efficiency will be nearly directly proportional to the crest value of the alternating plate voltage at all other times.

The maximum efficiency of a class B amplifier was shown to be about 65 per cent (page 355). Evidently this maximum efficiency is realizable only at the crest of the modulation cycle, at time t_5, Fig. 10-34*b*. With carrier alone, the alternating plate voltage will be reduced by one-half, as during the interval t_1t_2 (Fig. 10-34*b*), giving an efficiency of about 33 per cent. This is approximately the *average* efficiency of a class B linear amplifier, since the efficiency at time t_7 is evidently zero and the efficiency therefore varies between 0 and 65 per cent throughout the modulation cycle t_3 to t_8. Thus a class B amplifier, when used to amplify modulated waves, is relatively inefficient as compared with a class C amplifier.

Grid bias for class B amplifiers is normally obtained from a fixed source, since it must be maintained regardless of the presence or absence of excitation voltage. However, Wagener has shown that improved linearity may be obtained with certain tubes by permitting the bias to vary slightly with the modulation.[1] This may be done by supplying the major portion of the bias from a fixed source and the balance by means of a grid-leak resistor and suitable by-pass condenser; the bias will then vary with grid current and, therefore, with the modulation envelope. Tests for linearity may be conducted by plotting the characteristic curve of Fig. 10-11 or, better, by the use of a cathode-ray oscillograph which may be used under actual operating conditions as described on page 525 (Chap. 13).

Doherty High-efficiency Amplifier. The low efficiency of a class B linear amplifier has inspired the development of more efficient methods of amplifying modulated waves. One of the best known of these is due to Doherty.[2] The Doherty amplifier uses two tubes, one operating as a conventional class B amplifier

[1] W. G. Wagener, Simplified Methods for Computing Performance of Transmitting Tubes, *Proc. IRE*, 25, p. 47, January, 1937.

[2] W. H. Doherty, A New High Efficiency Power Amplifier for Modulated Waves, *Proc. IRE*, 24, p. 1163, September, 1936.

except that it is designed to operate at maximum efficiency with a grid excitation equal to the unmodulated carrier. Thus, with no modulation, E_{pm} is nearly equal to E_b, and this tube is incapable of producing, without distortion, the increased output required at the modulation crests. The second tube is operated with a bias such that its output is just zero with carrier alone impressed but operates with maximum efficiency at the crest of the modulation cycle. The first tube handles the entire output when the impressed wave has an amplitude less than the carrier, as from t_2 to t_4, Fig. 10-34*b*, and delivers carrier voltage only when the impressed wave exceeds the carrier, as from t_4 to t_6. The output of the second tube is zero during the time interval t_2 to t_4, but during the interval t_4 to t_6 it handles all power in excess of that de-

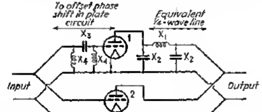


FIG. 10-35. Basic circuit of the Doherty high-efficiency amplifier.

livered by the first. This means that the first tube operates at maximum efficiency for one-half of the cycle (t_4 to t_6) while delivering high output, and operates at reduced efficiency only during the other half of the cycle (t_2 to t_4) when it is delivering low-power output. The second tube operates during one-half of the modulation cycle only; but, as will be shown shortly, the crest value of its alternating plate voltage during this time is never much less than half the direct voltage, and its average efficiency is, therefore, comparatively high. Such a combination permits operation at an efficiency of from 60 to 65 per cent, a marked improvement over the 30 to 35 per cent efficiency of the conventional class B linear amplifier.

The circuit of Fig. 10-35 illustrates the principle of operation of the Doherty amplifier, the direct voltages being omitted to simplify the circuit. Tube 1 is connected to the output circuit through a

reactance network which is the equivalent of a quarter-wave-length transmission line. Such a line presents an impedance at its sending terminals equal to Z_0^2/Z_r, where Z_0 is the characteristic impedance of the line and Z_r is the terminating impedance, in this case the impedance of the load. The impedance presented to tube 1 by this line, with tube 2 removed, should be such as to produce the desired carrier output at maximum efficiency. Throughout the time interval t_2 to t_4, Fig. 10-34b, this impedance will remain constant, and the tube will produce a voltage across the load proportional to the impressed grid voltage (Fig. 10-36a) (where the time intervals correspond to those of Fig. 10-34).

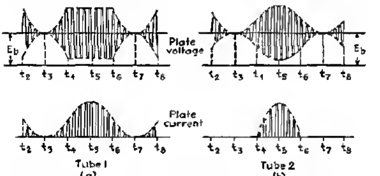


FIG. 10-36. Plate currents and voltages in the two tubes of a Doherty amplifier.

During the interval t_4 to t_6, however, tube 2 will produce an output, causing an increased current flow through the load. This means that for a given power output to the load, tube 1 will be called on for less power than if tube 2 were not in operation; *i.e.*, the apparent load resistance across the output terminals of the quarter-wave line is increased; and, therefore, the load impedance presented by the line to the plate circuit of tube 1 is decreased. This enables tube 1 to deliver an increased power output, owing to increasing plate current, without an increase in its alternating plate voltage, as shown in Fig. 10-36a. The crest value of the plate voltage E_{pm} is nearly equal to E_b from t_4 to t_6, resulting in a very low $e_{b\,min}$. As the tube delivers most of its power output during

this interval, it is readily seen that the average efficiency will be high.

The plate voltage on tube 2 is of course the voltage developed across the load and therefore varies as shown in Fig. 10-36*b*. It is comparatively high throughout the entire time of current flow through this tube, *i.e.*, $e_{b\ min}$ is low. Therefore, this tube's efficiency is also high.

It is interesting to note, for comparison purposes, that the plate voltage in a conventional class B linear amplifier is like that of tube 2 (Fig. 10-36*b*), whereas the plate current is like that of tube 1 (Fig. 10-36*a*), and thus a much larger portion of the plate-current flow takes place at higher instantaneous plate voltages. Herein lies the basic reason for the improvement in efficiency effected by the Doherty amplifier.

The equivalent quarter-wave line used in the plate circuit of tube 1 for impedance inversion causes a phase shift of 90 deg. It is, therefore, necessary to correct for this shift before combining the outputs of the two tubes. This is done by means of another quarter-wave network in the grid circuit of tube 1 using reactances of the opposite sign to those used in the plate circuit.

The networks of Fig. 10-35 are easily designed. In the plate circuit, quarter-wave conditions will be realized by making $X_1 = X_2 = Z_0$, where Z_0 is the characteristic impedance of the line. If R_1 is the correct load resistance to be used in the plate circuit of tube 1, the actual load presented by the output circuit should be $R_1/4$ and the characteristic impedance of the line Z_0 should be $R_1/2$. This combination will present a normal impedance of R_1 at the plate of tube 1 when tube 2 is inactive.

The design of the grid circuit network is carried out in a similar manner.

Applications of Class B Amplifiers. As will be more fully presented in Chap. 13, modulation may be applied to radio transmitters at the final amplifier, whence class C amplifiers are used throughout in the amplification of the r-f signal; or modulation may be applied at an earlier (lower power level) stage, after which class B amplifiers are used for further amplification. The former procedure seems to be the more popular and class B amplifiers are, therefore, not so widely used as class C. However the Doherty high-efficiency amplifier continues to be used in most applications where low-level modulation is applied.

Other high-efficiency systems than Doherty's have been suggested but no attempt will be made to describe them here.[1]

Grounded-grid Amplifier. Triode tubes may be used as r-f amplifiers without neutralization if the grid is grounded (instead of the cathode as in the conventional amplifier).[2] A circuit of such an amplifier is shown in Fig. 10-37. The input circuit may consist of the tuned output circuit of the preceding amplifier, but the type of input circuit is unimportant except that a d-c path must be provided between cathode and ground. No source of grid bias is shown, but any of the standard methods may be used.

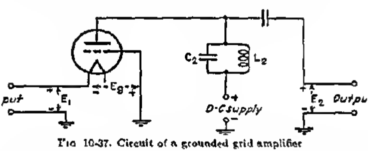


FIG. 10-37. Circuit of a grounded grid amplifier

Since the grid is at ground potential, any current flowing through the plate-to-grid capacitance will return directly to the plate circuit through the ground instead of returning through the input circuit as in the conventional amplifier. Thus the control grid serves in a dual capacity as a control and screen grid, and no neutralization is normally needed.

The gain of this amplifier may be readily found by first drawing the equivalent circuit as in Fig. 10-38. Here the plate-cathode circuit of the tube is replaced by the equivalent generator $-\mu E_g$ and the internal resistance r_p. The resistance R_L represents the resonant impedance of L_2 and C_2 in parallel. The circuit is seen

[1] One of these breaks the plate-current pulse into three sections, each of which has a nearly constant amplitude. Each section is then amplified by a separate tube operating at relatively high efficiency. For details see Sidney T. Fisher, A New Method of Amplifying with High Efficiency a Carrier Wave Modulated in Amplitude by a Voice Wave, *Proc. IRE*, 34, pp 3P-13P, January, 1946.

[2] M. C. Jones, Grounded-grid Radio-frequency Voltage Amplifiers, *Proc IRE*, 32, pp. 423-429, July 1944.

to be the same as that of Fig. 3-24, page 64, with the exception of the added input voltage $\mathbf{E}_1$ which Fig. 10-37 shows to be in series with the plate circuit.

From Fig. 10-38 we may write

$$\mathbf{E}_1 + (-\mu \mathbf{E}_g) = \mathbf{I}_p(r_p + R_L) \qquad (10\text{-}64)$$

From Fig. 10-37 it is seen that

$$\mathbf{E}_g = -\mathbf{E}_1 \qquad (10\text{-}65)$$

and putting this relation into Eq. (10-64) gives

$$(1 + \mu)\mathbf{E}_1 = \mathbf{I}_p(r_p + R_L) \qquad (10\text{-}66)$$

or

$$\mathbf{E}_1 = \mathbf{I}_p \frac{r_p + R_L}{1 + \mu} \qquad (10\text{-}67)$$

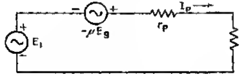


FIG. 10-38. Equivalent circuit of Fig. 10-37.

The output voltage $\mathbf{E}_2$ is evidently

$$\mathbf{E}_2 = \mathbf{I}_p R_L \qquad (10\text{-}68)$$

The gain may now be found by dividing Eq. (10-68) by Eq. (10-67) to give

$$\mathbf{A} = \frac{\mathbf{E}_2}{\mathbf{E}_1} = \frac{(1 + \mu)R_L}{r_p + R_L} \qquad (10\text{-}69)^*$$

It is of interest to compare this equation with a similar one for the conventional (grounded-cathode) type of amplifier. The output voltage for the conventional amplifier is equal to the voltage set up across the load impedance by the plate current, therefore Eq. (10-68) also applies to a conventional amplifier. The input voltage, however, is equal to the grid voltage, which may be found from Eq. (3-20), page 63. Dividing E_2 by E_g gives

$$\mathbf{A} = \frac{-\mu R_L}{r_p + R_L} \qquad (10\text{-}70)^*$$

which is seen to differ from Eq. (10-69) only by the numeric 1 which is added to μ in the latter equation, and by the negative sign. Thus a grounded-grid amplifier produces the same gain as a conventional amplifier with a tube having an amplification factor of $(1 + \mu)$ but with no phase reversal.

The input impedance of a grounded-grid amplifier is relatively low. From Fig. 10-37 it is evident that the current flowing in the input circuit is equal to the plate current of the tube. We may therefore write for $\mathbf{Z}_1$, the input impedance,

$$\mathbf{Z}_1 = \frac{\mathbf{E}_1}{\mathbf{I}_p} = \frac{r_p + R_L}{1 + \mu} \qquad (10\text{-}71)^*$$

the substitution for $\mathbf{E}_1$ being made from Eq. (10-67)

From the foregoing it is evident that the grounded-grid amplifier is characterized by having equal output and input currents, and by a low input impedance and a high output impedance. It is of interest to compare this with the cathode-follower amplifier (described on page 371) in which the input and output *voltages* are virtually equal and in which the input impedance is *high* and the output impedance *low*. The cathode-follower amplifier was described as an a-f amplifier, but it may be used equally as well at radio frequencies by replacing the load resistance with a resonant circuit

At very high frequencies some neutralization may be required with the grounded-grid amplifier, just as with a conventional pentode amplifier, to balance out the effect of external capacitances and any plate-cathode capacitance not eliminated by the presence of the grounded grid. However the small amount of neutralization required makes the problem of securing satisfactory balance very much simpler than in the conventional neutralized, grounded-cathode amplifier.[1]

Cathode-coupled Amplifier. For many applications the low input impedance of the grounded-grid amplifier is objectionable, but if a cathode-follower is used as a driver, the resulting two-stage amplifier presents both a high input impedance and a high output impedance, a lower noise level than a conventional pentode amplifier, and yet compares favorably with the pentode amplifier in

[1] For a further discussion of these problems, see C E Strong, The Inverted Amplifier, *Electronics*, 13, p. 14, July, 1940.

gain and stability. The circuit is shown in Fig. 10-39, the two tubes being commonly contained in a single envelope with a common cathode as indicated.[1] Tube 1 is the cathode follower with its output circuit, L_1C_1, in series with the cathode. The grounded-grid amplifier, tube 2, receives its excitation from the L_1C_1 circuit and delivers power into the output tank circuit L_2C_2. Neither tube requires neutralization except at very high frequencies. In

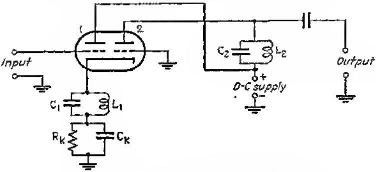


FIG. 10-39. Cathode-coupled amplifier.

the cathode follower, feedback exists through the grid-cathode capacitance, but the output and input voltages are so nearly equal that the feed-back current is relatively small.

The cathode resistor R_k and by-pass condenser C_k provide bias for both tubes.

Problems

10-1. A certain r-f voltage amplifier, using the circuit of Fig. 10-1, has the following constants: $Q = 150$, $L = 3.7\ \mu h$, $R_g = 0.25$ megohm, $g_m = 1600$ μmhos, $r_p = 1.8$ megohms. The frequency of resonance is 20 Mc. Compute (a) the gain at resonance and (b) the band width.

10-2. A certain r-f voltage amplifier is to be designed for operation at 2.5 Mc using the circuit of Fig. 10-1. The gain is to be 90 (39 db) per stage and the band width is to be 20 kc. $R_g = 0.5$ megohm. The coefficients of the tube to be used are $g_m = 1500\ \mu$mhos, $r_p = 1.7$ megohms. Find (a) the inductance of the coil, (b) the Q of the coil to produce the required gain and band width, and (c) the capacitance of the tuning condenser at resonance.

10-3. A r-f amplifier, using the circuit of Fig. 10-3, has the following con-

[1] For further information, see G. C. Sziklai and A. C. Schroeder, Cathode-coupled Wideband Amplifiers, *Proc. IRE*, **33**, pp. 701–709, October, 1945; and Keats A. Pullen, The Cathode-coupled Amplifier, *Proc. IRE*, **34**, pp. 402–405, June, 1946.

stants. $L_1 = 6\ \mu h$, $L_2 = 5\ \mu h$, coupling between L_1 and L_2 = 40 per cent, Q (of coil L_2) = 170, g_m = 1200 μmhos, r_p = 1.4 megohms, frequency of resonance = 20 Mc. Compute (a) the gain at resonance, (b) the band width, and (c) the capacitance of the condenser C_2 at resonance.

10·4. Determine L and C in the tank circuit of a class C amplifier using a triode under the following conditions: E_b = 2000 volts, E_c = −160 volts, E_p = 1280 volts, power delivered to the tank circuit = 375 watts, f = 2500 kc. Carry out the design for (a) Q = 12.5 and (b) Q = 75.

10·5. Determine the percentage of harmonic in the plate voltage for the two parts of Prob. 10-4, assuming that the second harmonic in the plate current is 60 per cent of the fundamental.

10·6. The tank current in a certain class C amplifier is 0.5 amp. The reactance of the tank condenser is 1280 ohms. E_b = 1000 volts, and E_c = −100 volts. (a) What is $e_{b\,min}$? (b) If the amplifier is operating under conditions of maximum efficiency, what is E_g?

10·7. Carry out the approximate solution of the tank circuit of a class C amplifier as given on page 421. E_b = 1000 volts, P_o = 50 watts, Q = 15, f = 42 Mc. Solve for L and C of the tank circuit.

Design Problems[1]

10·8. Carry out the approximate design of a class C amplifier, using the 833 tube, to give approximately the maximum possible output at a direct plate voltage of 2500. $\theta_p = 75°$.

10·9. Carry out the approximate design of a class C amplifier using a triode tube, the characteristics of which are secured from a tube manual.

10·10. Repeat Prob. 10-9 for a tetrode.

10·11. Repeat Prob. 10-9 for a pentode.

[1] See footnote on p. 215.

CHAPTER 11

OSCILLATORS

In Chaps. 9 and 10 it was seen that an amplifier has the property of producing a-c power output from a d-c power source in the plate circuit when its grid is under control of another source of a-c power. The amount of power thus produced in the output circuit is many times that required to excite the grid so that, in effect, the impressed a-c power is amplified, although the output power is actually converted from the local source of direct current.

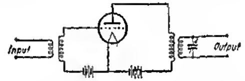


FIG. 11-1. Circuit of a simple r-f amplifier.

Since the power available in the output is so much greater than that required in the input circuit, it should be possible to divert a portion of this power back to the input and thus make the amplifier self-excited. An amplifier so connected and producing an output is known as an *oscillator*.

Figure 11-1 shows a simple amplifier circuit with a resonant type of load. In Fig. 11-2 is shown the same circuit but with the input leads connected across the output terminals. The input leads are shown crossed, to compensate for the 180-deg phase reversal in the tube. The amount of power that may be drawn from the output terminals is now less than in the circuit of Fig. 11-1 by the amount required for excitation of the grid, but in a properly designed circuit this difference is very small.

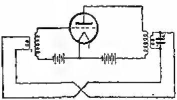


FIG. 11-2. Circuit of Fig. 11-1 with the input circuit excited from the output of the same tube.

An amplifier with a resonant circuit load was chosen for this example purposely. If an amplifier with a pure resistance load is used and an attempt is made to generate oscillations in the manner of Fig. 11-2, the result will be a failure. The reason for this is to be found in the nonuniform flow of power through the plate circuit. The plate current of any amplifier is pulsating, or at least fluctuating between wide limits, so that the power flow to the load is not constant. There are periods during which the power delivered is insufficient to supply the losses, and the oscillations, even if started in some manner, would immediately die out. The power flow to the output circuit may actually be negative during a small portion of the cycle in high-efficiency oscillators, just as in the class C amplifier. In fact, for high efficiency, the oscillator should operate in the same manner as the class C amplifier, having the same general wave shapes of currents and voltages as shown in Fig. 10-15 or 10-17. It is therefore necessary that energy storage be supplied to carry the load through the periods of low or negative power supply from the tube. The resonant circuit supplies this storage in the fields of its condenser and coil.

As in the class C amplifier this need for energy storage in a vacuum-tube oscillator may be likened to the need for mechanical energy storage in the case of a reciprocating engine. The energy is supplied by such an engine throughout only a portion of the complete cycle, and the flywheel must store sufficient energy during this period to carry the load until the next pulse of energy is received from the piston.

The energy storage for an oscillator may be supplied in other ways than by an electrically resonant circuit, as in the case of the piezoelectric and multivibrator oscillators described later in this chapter, but the energy must be stored in some manner, or oscillations cannot be sustained. The resonant circuit generally provides the simplest and most convenient method for providing this storage, besides ensuring a nearly sinusoidal output wave.

Oscillator Circuits. Any amplifier circuit that includes provision for energy storage may be used for the generation of oscillations, but certain modifications may be made that result in simpler construction or more efficient operation as an oscillator. One of the most common oscillator circuits is the Hartley, Fig. 11-3. It is seen to be fundamentally the same as the circuit of Fig. 11-2 except that the grid excitation is obtained directly from the reso-

nant circuit instead of through a transformer. There may or may not be mutual induction between coils L_1 and L_2. The d-c sources for grid and plate (shown as batteries in Fig. 11-3) are necessarily placed between the resonant circuit and their respective tube elements, as the return to the cathode is common for both the plate and the grid circuits.

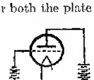


FIG. 11-3. Basic Hartley oscillator circuit.

The Hartley circuit cannot be used in the exact form of Fig. 11-3 except for very low power oscillators. The power supply for the plate and grid voltages is usually obtained from rectifiers rather than from batteries, and the filament-heating power is obtained from the secondary of a transformer or from a generator. Such power-supply units are generally grounded so that all have one common terminal; but, even if not, the capacitance between transformer windings, etc., is sufficient virtually to short-circuit the coils L_1 and L_2 at all but the very lowest frequencies.

This difficulty may be avoided by inserting the plate and grid d-c power supplies at points b and c, respectively, so that no alternating difference in potential exists between them. This scheme is sometimes employed for low-voltage tubes. The principal objection is that the coils L_1 and L_2 must be insulated from ground by an amount equal to the crest value of the alternating voltage induced in them plus the direct voltage, and the condenser C has a maximum voltage impressed across it equal to the total crest value of the alternating voltage of both coils plus the total direct voltage. With high-voltage tubes this seriously increases the cost of insulation, especially as the condenser C is frequently an air-insulated condenser which, in case of a flashover, will restore itself automatically if only the alternating voltage of the oscillator is impressed across it; whereas, with the plate power supply at b, Fig. 11-3, a power arc from the d-c source may follow with consequent danger to equipment.

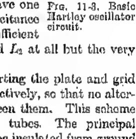


FIG. 11-4. Hartley oscillator circuit; parallel feed.

The objections raised in the preceding paragraph are entirely

eliminated by the circuit of Fig. 11-4. The d-c plate supply is fed to the tube through an inductance L_x, the reactance of which is high compared with the impedance of the resonant circuit. A condenser C_p prevents the direct current from being short-circuited through coil L_2, while permitting free passage of the alternating component of plate current. Grid bias is supplied by the drop through resistance R_g set up by the direct component of grid current. Condenser C_g by-passes the alternating components so that the voltage across the resistance is constant throughout the cycle, even though the instantaneous grid current varies. The use of a grid leak and condenser in place of a battery or other fixed-voltage source, to secure grid bias, is essential if the oscillator is to be self-starting. For efficient operation an oscillator should be biased below cutoff, like a class C amplifier; and if full bias is applied at the same time as the plate voltage, the plate current will be zero and there will be no power input to start the oscillations. With a grid leak the bias will be zero when the plate voltage is applied, plate current will flow, and oscillations will start immediately, automatically building up the bias to its correct value.

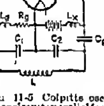


FIG. 11-5 Colpitts oscillator circuit; parallel feed

Another commonly used circuit is the Colpitts, Fig. 11-5. Its chief point of difference from the Hartley is in the method of securing the grid and plate voltages from the tank circuit, by tapping the condenser (placing two in series so that their series capacitance is equal to the total desired) instead of the inductance. The grid is completely isolated from both cathode and anode by condensers C_1, C_2, and C_p, and the grid leak must be placed directly between the grid and cathode to secure a return circuit for the direct component of grid current, the condenser C_1 serving double duty as grid condenser and as part of the tank capacitance. The choke L_g prevents absorption of r-f power by the grid leak but is often omitted, if R_g is not too low, as it has been found that the increased load imposed by the grid leak tends to stabilize the frequency. This is due, in part at least, to the more constant resistance between grid and cathode that the grid leak provides. The input resistance of the tube alone varies considerably through-

out the cycle of impressed voltage, so placing a constant resistance of lower value in parallel produces a net resistance that is relatively independent of the instantaneous grid voltage.

Many modifications of these two circuits may be used such as the circuit of Fig. 11-6. Here the excitation for the grid is obtained through the mutual inductance between coils L_1 and L_2.

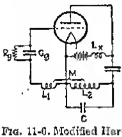


FIG. 11-6. Modified Hartley circuit.

The tuned-grid, tuned-plate circuit of Fig. 11-7 has found great favor among designers of higher frequency oscillators. There is no mutual inductance between coils L_1 and L_2, the energy being fed back to the grid circuit through the capacitance between the plate and grid of the tube. The operation of the circuit may be made more clear by rearranging it in the form of Fig. 11-8 where there is no common field between the two coils L_1 and L_2. C_{gp} represents the capacitance between the grid and plate inside the tube and serves the same function as the mutual inductance between the two coils of Fig. 11-6. Actually C_{gp} should be shown connected directly between the grid and plate terminals of the tube instead of across the resonant circuit, but the two condensers C_g and C_p which act in series with C_{gp} in passing the feed-back current are so very much larger that they may be neglected, and the operation can be depicted with equal accuracy and greater clarity as shown.

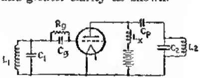


FIG. 11-7. Tuned-grid, tuned-plate oscillator.

General Equations of an Oscillatory Circuit.[1] All tuned-circuit oscillators may be represented by the basic circuit of Fig. 11-9. Applying this circuit to the Hartley oscillator of Fig. 11-4, for example, L and C represent the tank circuit, L being equal to L_2, and C having a reactance equal to the total reactance of L_1 and C

[1] This section may be omitted without loss of continuity except that the analysis of the R-C oscillator (p. 462) is based on this material.

of Fig 11-4. The resistance R represents the power consumed by the tank circuit in useful output and losses while r represents the plate-to-cathode circuit of the tube of Fig. 11-4. The latter resistance must necessarily be negative since the tube is supplying power to the circuit, as will be further demonstrated in the following development. The grid-to-cathode circuit is not represented in Fig 11-9, the grid losses being included in R.

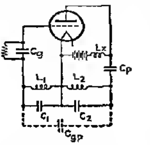


FIG. 11-8 Equivalent circuit of Fig 11-7.

Let us assume that the three currents in Fig 11-9 are i_1, i_2, and i_3 and that the assumed positive direction of flow of each is as indicated by arrows We may

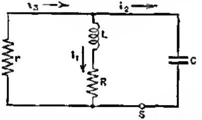


FIG 11-9 Equivalent circuit of an oscillator.

then write the three Kirchhoff law equations for this circuit as follows

$$i_1 + i_2 - i_3 = 0 \tag{11-1}$$

$$i_3 r + L\frac{di_1}{dt} + i_1 R = 0 \tag{11-2}$$

$$i_1 R + L\frac{di_1}{dt} - \frac{q}{C} = 0 \tag{11-3}$$

where q is the instantaneous charge on condenser C. Since $i_2 = dq/dt$ we may eliminate q from Eq. (11-3) by differentiating and substituting for dq/dt to give

$$R\frac{di_1}{dt} + L\frac{d^2i_1}{dt^2} - \frac{i_2}{C} = 0 \tag{11-4}$$

We may now solve Eqs. (11-1), (11-2), and (11-4) simultaneously for the three currents. As a first step solve Eq. (11-1) for i_3 and substitute this value in Eq. (11-2) to give

$$i_1r + i_2r + L\frac{di_1}{dt} + i_1R = 0 \tag{11-5}$$

Equation (11-5) may now be solved for i_2 and this value then substituted in Eq. (11-4). When the various terms have been gathered together, the result is

$$L\frac{d^2i_1}{dt^2} + \left(R + \frac{L}{rC}\right)\frac{di_1}{dt} + \left(\frac{R}{rC} + \frac{1}{C}\right)i_1 = 0 \tag{11-6}$$

This is a linear differential equation of the second order with constant coefficients. A method of solving such an equation is to assume its solution from inspection of the equation and then substitute this presumed solution back into the original equation for verification or rejection. As it turns out, the solution of this type is of the general form

$$i_1 = A\epsilon^{mt} \tag{11-7}$$

where A is a constant of integration and m must be evaluated to satisfy Eq. (11-6). To check the validity of this solution and to evaluate m, Eq. (11-7) must be substituted into Eq. (11-6) to give

$$Lm^2A\epsilon^{mt} + \frac{RrC + L}{rC}mA\epsilon^{mt} + \frac{r + R}{rC}A\epsilon^{mt} = 0 \tag{11-8}$$

Canceling $A\epsilon^{mt}$ out of this equation gives

$$Lm^2 + \frac{RrC + L}{rC}m + \frac{r + R}{rC} = 0 \tag{11-9}$$

This is a quadratic equation in the variable m and may be solved in the usual manner to give[1]

[1] Equation (11-9) is of the general form $ax^2 + bx + c = 0$, the solution of which is

$$m = -\frac{RrC + L}{2rCL} \pm \sqrt{\left(\frac{RrC + L}{2rCL}\right)^2 - \frac{r + R}{rCL}} \quad (11\text{-}10)$$

Equations (11-9) and (11-10) show that Eq. (11-7) is a solution of Eq. (11-6) provided m satisfies Eq. (11-10). Equation (11-10) also shows that there are two different values of m that will satisfy the equation. Under such circumstances a more general solution may be obtained by taking the sum of the two solutions or

$$i_1 = A_1\epsilon^{m_1t} + A_2\epsilon^{m_2t} \quad (11\text{-}11)$$

where A_1 and A_2 are integration constants and m_1 and m_2 are defined by Eq. (11-10), using the plus sign for m_1 and the minus sign for m_2. [The student should prove for himself that Eq. (11-11) is a solution of Eq. (11-6) by substituting it back into Eq. (11-6) using the values of m_1 and m_2 from Eq. (11-10).]

To simplify the notation, let

$$m_1 = \alpha + \beta \quad (11\text{-}12)$$

$$m_2 = \alpha - \beta \quad (11\text{-}13)$$

where

$$\alpha = -\frac{RrC + L}{2rCL} = -\left(\frac{R}{2L} + \frac{1}{2rC}\right) \quad (11\text{-}14)$$

$$\beta = \sqrt{\left(\frac{RrC + L}{2rCL}\right)^2 - \frac{r + R}{rCL}} \quad (11\text{-}15)$$

We are now ready to solve for the two integration constants. These may be found by substituting into Eq. (11-11) the coordinates of two known points on the curve of i_1 vs. t or by substituting the coordinates of one known point together with the known derivative at that point. The latter method is the one which must normally be employed and will be used here. To perform this step, it is helpful to assume that a switch has been inserted at S, Fig. 11-9, which is now open, and that the condenser has been charged to a voltage E. This is not the way in which the circuit is originally energized in an actual oscillator, but since we are interested only in the steady-state condition, the method of initiating the current flow is unimportant.

$$x = -\frac{b}{2a} \pm \sqrt{\left(\frac{b}{2a}\right)^2 - \frac{c}{a}}$$

Let the switch be closed at the time $t = 0$. Since the current i_1 was zero prior to closing the switch, it must still be zero at the instant of closing the switch. This is shown by Eq. (11-3) since if i_1 jumped immediately to some value other than zero, the derivative di_1/dt would be infinite and the equation would not be satisfied. If, on the other hand, the current is still zero at $t = 0$ but immediately starts to rise, Eq. (11-3) becomes

$$L\frac{di_1}{dt} - \frac{q}{C} = 0 \qquad (\text{at } t = 0) \tag{11-16}$$

But q/C at $t = 0$ is equal to the voltage E impressed across the condenser, and Eq. (11-16) becomes

$$L\frac{di_1}{dt} - E = 0 \qquad (\text{at } t = 0) \tag{11-17}$$

We are now in a position to say that the coordinates of our known point are 0,0, *i.e.*, when t equals zero, i_1 is also zero. Furthermore the slope of the curve at this point is given by Eq. (11-17), so that $di_1/dt = E/L$ at $t = 0$. We may now substitute the coordinates of the known point into Eq. (11-11) and the value of the derivative into the derivative of Eq. (11-11), which is obviously

$$\frac{di_1}{dt} = m_1A_1\epsilon^{m_1t} + m_2A_2\epsilon^{m_2t} \tag{11-18}$$

Substitution of the coordinates into Eq. (11-11) gives

$$0 = A_1 + A_2 \tag{11-19}$$

Substitution of the derivative at $t = 0$ and Eqs. (11-12) and (11-13) into Eq. (11-18) gives

$$\frac{E}{L} = (\alpha + \beta)A_1 + (\alpha - \beta)A_2 \tag{11-20}$$

These two equations may be solved simultaneously for A_1 and A_2 to give

$$A_1 = \frac{E}{2\beta L} \tag{11-21}$$

$$A_2 = -\frac{E}{2\beta L} \tag{11-22}$$

Thus the final equation for i_1 is found by substituting Eqs. (11-12), (11-13), (11-21), and (11-22) into (11-11) giving

$$i_1 = \frac{E}{\beta L}\epsilon^{\alpha t}\left(\frac{\epsilon^{\beta t} - \epsilon^{-\beta t}}{2}\right) \tag{11-23}$$

The parenthetical term is the hyperbolic sine of βt which varies from zero to infinity in a rising curve as t is increased from zero, apparently indicating that the circuit will not oscillate since there is no periodic variation of the current. But inspection of Eq. (11-15) shows that β may be either real or imaginary, depending on the relative size of the two terms under the radical. If β is imaginary, it is better to rewrite Eq. (11-15) as

$$\beta = j\beta' = j\sqrt{\frac{r + R}{rCL} - \left(\frac{RrC + L}{2rCL}\right)^2} \tag{11-24}$$

where β' is a real number. Replacing β with $j\beta'$ in Eq. (11-23) gives

$$i_1 = \frac{E}{j\beta' L}\epsilon^{\alpha t}\left(\frac{\epsilon^{j\beta' t} - \epsilon^{-j\beta' t}}{2}\right) \tag{11-25}$$

If the j in the denominator is placed inside the parentheses, the parenthetical term becomes the sine of $\beta' t$ and we may therefore rewrite Eq. (11-25) as

$$i_1 = \frac{E}{\beta' L}\epsilon^{\alpha t}\sin\beta' t \tag{11-26}$$

But the usual symbol for the angle of a sine wave is ωt; therefore it is preferable to replace β' in Eq. (11-26) with ω to give

$$i_1 = \frac{E}{\omega L}\epsilon^{\alpha t}\sin\omega t \tag{11-27}^*$$

Equation (11-27) shows that, if β is imaginary, the current flowing through the inductance of Fig. 11-9 will be periodic and the circuit will oscillate, but that the amplitude is normally either increasing or decreasing, depending on whether α is positive or negative. Oscillators should be so designed that α is positive as the oscillations start, causing the amplitude to increase, but α will decrease to zero as the amplitude rises, owing to an increase in r_p. The plate resistance r_p increases because, as the current amplitude rises, the grid voltage will swing farther and farther

below cutoff on its negative half cycles, where r_p is essentially infinite, and the average value of r_p over a cycle will rise.

The plate resistance r_p does not appear in Eq. (11-14) but r is directly affected by r_p. This may be shown by noting that r, in the oscillators thus far considered, is the resistance seen looking back into the plate-cathode terminals and is, therefore,

$$r = -\frac{E_p}{I_p} \tag{11-28}$$

(The minus sign must be used since the positive polarity of I_p was originally assumed on the basis of power flowing *out* of the plate circuit whereas r must be measured looking *into* the plate circuit.) The plate voltage E_p is nearly constant regardless of reasonable changes in the circuit or load. This was demonstrated for class C amplifiers on page 422, and very similar conditions normally prevail in an oscillator. Thus any increase in r_p must result in a decrease in I_p and therefore an increase in r. This will cause the second term on the right-hand side of Eq. (11-14) to decrease until it is equal to the first, and α will be equal to zero.

After the oscillations have reached the steady-state condition and α is therefore zero, the equation for ω may be simplified by noting that the second term under the radical in Eq. (11-24) is equal to α^2 and is therefore zero. We may then rewrite Eq. (11-24) as

$$\omega = \sqrt{\frac{1}{LC}\left(\frac{r+R}{r}\right)} \tag{11-29}$$

Since α is zero, we may solve for r from Eq. (11-14) to obtain

$$r = -\frac{L}{RC} \tag{11-30}$$

When this relation is substituted into Eq. (11-29) the result is

$$\omega = \sqrt{\frac{1}{LC} - \frac{R^2}{L^2}} \tag{11-31}^*$$

If both sides of this equation are squared and both sides of the resulting equation are multiplied by $1/\omega^2$, we obtain

$$1 = \frac{1}{\omega^2 LC} - \frac{R^2}{\omega^2 L^2} \tag{11-32}$$

Substituting $1/Q^2$ for R^2/ω^2L^2, this equation may be solved for ω to give

$$\omega = \sqrt{\frac{Q^2}{1+Q^2}}\sqrt{\frac{1}{LC}} \tag{11-33}^*$$

We have already seen, in our study of class C amplifiers, that Q should be at least 10 or 12 to secure good wave shape and stability, thus it is evident that $Q^2/(1 + Q^2)$ is virtually equal to unity. The frequency of oscillation is then, for all practical purposes,

$$f = \frac{1}{2\pi\sqrt{LC}} \tag{11-34}^*$$

From the foregoing discussion we may conclude that the current flowing in the inductance of the tank circuit of an oscillator under steady-state conditions is

$$i_1 = \frac{E}{\omega L}\sin \omega t \tag{11-35}^*$$

which is Eq. (11-27) with $\alpha = 0$. The term $E/\omega L$ is evidently the crest value of the current; therefore, the voltage E, to which the condenser was assumed to be originally charged, must be the crest value of the voltage across the tank inductance L, which, as was shown in the preceding chapter, is very nearly equal to the direct plate voltage E_b of the tube.

Applications of the General Equations. The foregoing general equations will be used later to analyze R-C oscillators. They are of little value in the actual design of tuned-circuit oscillators but should assist the student in envisioning their performance under various conditions. As an example, these equations are helpful in determining why an oscillator may fail to oscillate. From Eq. (11-23) we saw that for the current to be a sustained sine wave, α had to be zero (or positive) and β had to be imaginary. Evidently, then, two reasons for failure to oscillate are (1) that α is negative or (2) that β is real.

The proper remedies to apply for correcting a condition where β is real may be found from Eq. (11-31) where ω will be imaginary, if β is real. An imaginary value of ω is evidently due to the second term under the radical of Eq. (11-31) exceeding the first, and a remedy is evidently to decrease the ratio R/L (increase Q).

This may be done by decreasing the required power output which will decrease R and so make ω real.

A generally better solution is to increase the Q of the circuit by increasing C. That this is an effective procedure may be seen by comparing the last term of Eq. (11-32) with Eq. (10-62), page 420, whence it is evident that

$$\frac{R^2}{\omega^2 L^2} = \left(\frac{1}{2\pi}\,\frac{\text{energy dissipated per cycle}}{\text{energy stored per cycle}}\right)^2 \qquad (11\text{-}36)^*$$

Since it was demonstrated in the discussion of class C amplifiers that an increase in C increases the stored energy for a given power output, it follows from Eq. (11-36) that an increase in C will result in a decrease in the last term of Eq. (11-32) and, therefore, in the second term under the radical of Eq. (11-31), provided the energy dissipated per cycle (power output) is maintained constant. Referring again to the discussion of the class C amplifier, an increase in C must be accompanied by a reduction in L to keep the frequency unchanged, but this must be accompanied by an even greater decrease in R if the power output is to be maintained constant.

Equation (11-14) shows that if α is to be positive, as required, at the time oscillations are to start, $1/2rC$ (which is the negative term) must be larger than $R/2L$. In the limit, as the oscillator approaches the steady-state condition, these two terms must be equal or

$$-\frac{1}{2rC} \geqq \frac{R}{2L} \qquad (11\text{-}37)$$

Since α is the attenuation factor, it is to be expected that its magnitude will be largely influenced by the amount of power flowing into and out of the tank circuit. This is borne out by the relation of Eq. (11-37) where it may be seen that a decrease in r, indicating an increase in the power delivered to the circuit, will tend to make α positive (or at least less negative) and a decrease in R, indicating less power being absorbed by the circuit, will have the same effect. Thus if an oscillator will not oscillate due to a negative value of α, the remedy is to increase the power input, as by increasing the grid excitation, or decrease the power demanded by the load.

Summarizing the foregoing it is evident that β will be imaginary,

as required, if Q is sufficiently large and therefore determines the ability of the circuit to oscillate *if adequately energized*, while α determines the adequacy of the power being supplied by the tube.

Power Oscillators. Most oscillators fall into one of two groups: *power oscillators* and *frequency-controlling oscillators.* The latter are probably the more common and are used primarily to hold the frequency of a given radio transmitter constant within very close limits. For the most part they are not required to generate much power but are used to drive class C amplifiers which produce the desired output.

Power oscillators are those which furnish the desired power output without the need of further amplification, and their frequency stability, although important, is secondary to their output requirements. They are operated with the same type of plate-current flow as class C amplifiers, and the same design procedure may be followed. The Q of the tank circuit determines the frequency stability as well as the wave shape, a high Q being necessary for good stability. Power oscillators are designed with a Q no higher than the required frequency stability makes absolutely necessary, since an increase in Q is accompanied by a decrease in efficiency. A commonly used minimum figure for Q is 12 which, as shown by Eq. (10-62), represents a ratio of energy stored per cycle to energy dissipated of about 2.

Conservative design procedure requires that Q be taken as the ratio of that portion of the tank circuit reactance which is coupled to the plate, to the total effective series resistance of the tank.[1] In the Hartley circuit of Fig. 11-4, for example, only the reactance of the coil L_1 should be used in determining Q; in the Colpitts circuit of Fig. 11-5 only the reactance of that portion of the coil L which resonates with the capacitance C_1 should be used; *i.e.*, Q is the ratio of the reactance of C_1 to the effective resistance. This makes the design exactly analogous to that of a class C amplifier where the tank circuit consists only of the plate inductance and capacitance, the grid being energized from a separate source; and the method outlined in Chap. 10, starting on page 414, may be used in the design of a power oscillator.

[1] The effective series resistance in an oscillator tank circuit is the same as R_t, Fig. 10-18a, in a class C amplifier.

Example. Let us assume that an oscillator is to be designed using an 833 tube with a direct plate voltage of 3000 volts. Reference to page 410 will show that this is the same tube and same plate voltage as were assumed for the class C amplifier example; therefore, the design procedure for determining the correct alternating voltages for the tube will follow that of the class C amplifier. Thus we find from the class C amplifier example that the tube will deliver 952 watts at 79.6 per cent efficiency for an assumed θ_p of 75 deg and will require alternating voltages of 2000 and 215 on the plate and grid, respectively.

The next step in the design is to determine the tank constants which will produce these voltages and output. Let us assume that the circuit is to be that of Fig. 11·4 and that L_x and C_p are sufficiently large so that their effect is negligible. It will also be helpful to rearrange the circuit until it looks more like the class C amplifier of Fig. 10-14. This has been done in Fig. 11-10, and the resistances R_1 and R_2 have been inserted to represent the power absorbed by the tank circuit, where $(R_1 + R_2)$ corresponds to R_L of the class C amplifier in Fig. 10-18. Thus R_1 and R_2 represent the loss in the coils L_1 and L_2 and the useful power output, and the relative magnitudes

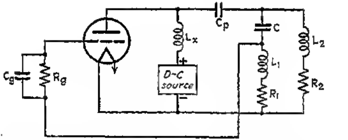


FIG. 11-10. Circuit of Fig. 11-4 rearranged to conform more closely to the class C amplifier circuit of Fig. 10-14.

of these two resistances depend upon the method of coupling the output circuit to the tank. Since their relative magnitudes are normally of no importance, we shall merely consider their total value $R_L = R_1 + R_2$.

Comparison of the tank circuit of Fig. 11-10 with that of Fig. 10-18a shows that the coil L_2 is comparable to coil L and that C and L_1 in series, which have a net condensive reactance, correspond to the condenser C. We may therefore apply Eq. (10-42) to the oscillator by substituting L_2 for L. Assuming that $Q = 12.5$, this gives

$$\omega_0 L_2 = \frac{(2000)^2}{952 \times 12.5} = 336 \text{ ohms}$$

exactly as in the class C amplifier design on page 421.

The reactance of L_1 may now be found by noting that $\omega_0 L_2/\omega_0 L_1 = E_p/E_g$. Since all these factors are known except $\omega_0 L_1$, we may solve for that term.

$$\omega_0 L_1 = \frac{E_g}{E_p}\omega_0 L_2 = \frac{215}{2000} \times 336 = 36 \text{ ohms}$$

The reactance of the condenser is obviously equal to

$$\frac{1}{\omega C} = 336 + 36 = 372 \text{ ohms}$$

As in the class C amplifier, the size of the choke L_c and of the blocking condenser C_g is not critical and, following the procedure of the class C amplifier example on page 421, we may solve for the inductances and capacitances for an assumed frequency of 1000 kc as follows

$$L_1 = \frac{36}{2\pi 10^6} \times 10^6 = 5.75\ \mu\text{h}$$

$$L_2 = \frac{336}{2\pi 10^6} \times 10^6 = 53.5\ \mu\text{h}$$

$$C = \frac{1}{2\pi 10^6 \times 372} \times 10^{12} = 427\ \mu\mu\text{f}$$ 427 83386

$$C_g \geq \frac{1}{2\pi 10^6 \times 400} \times 10^{12} \geq 400\ \mu\mu\text{f}$$ = 397.88735

$$L_c \geq \frac{40{,}000}{2\pi 10^6} \times 10^3 \geq 6.4 \text{ mh}$$ = 63661977

The grid leak resistance may be found from $R_g = E_c/I_c$, or

$$R_g = \frac{144}{0.077} = 1870 \text{ ohms}$$ 1870.1298

The approximate method of tank circuit design, as given for the class C amplifier on page 421, may also be applied to oscillators. The reactance then obtained from Eq. (10-42) is that of L_1 for the Hartley oscillator of Fig. 11-4, and that of the condenser C_2 for the Colpitts oscillator of Fig. 11-5. [For the Colpitts oscillator Eq. (10-42) should read $1/\omega C_2 = E_p^2/P_o Q$.] In the Hartley oscillator the coils L_1 and L_2 are usually a single coil as indicated in Fig. 11-4 which must be made somewhat larger than the computed value for L_2. The condenser C must be of suitable size to tune this coil to the desired frequency. The grid and plate leads should then be brought out from the tube to taps on the coil and final adjustments of alternating grid and plate voltages can be made experimentally after the oscillator is in operation.

Increasing the grid voltage of an oscillator by moving the grid tap farther from the cathode tap increases the bias as well as the alternating voltage, keeping the maximum positive grid voltage approximately constant. This results in higher plate efficiency with very little change in output but with appreciable increase in grid driving power. In general the grid tap should be so placed as to provide the maximum grid voltage consistent with a reasonable size of coil and without requiring excessive grid driving power.

Frequency Stability. The term *frequency stability* refers to the ability of the oscillator to maintain constant frequency under operating conditions. The frequency of oscillation is obviously affected by any reactances in the circuit in addition to those of the tank. Most of these are constant and thus do not affect the frequency stability; but the tube capacitances vary somewhat with the electrode potentials, and the reactance introduced by the load may change with any variation in the power demand.[1] Variations in the tube capacitances are of greatest importance at very high frequencies where these reactances are of the same general order of magnitude as the components of the tank circuit.

Frequency stability is also affected by changes in the resistive components of the circuit, principally the effective load resistance and the plate resistance of the tube. It is also affected by the gain of the tube and therefore varies with any change in the transconductance. Thus it may be said that good frequency stability depends upon supplying the tube with d-c power supplies of good regulation and high degree of constancy to maintain μ, g_m, and r_p constant, and upon maintaining a constant load.

The effect on the frequency stability of changes in resistance may be seen by constructing a vector diagram of the fundamental components of the voltages and currents in the oscillator. To illustrate, the Hartley circuit will be used, the a-c circuit elements of which are shown in Fig. 11-11*a* where the effect of the blocking condenser, plate choke, and grid current are considered negligible and where R_1 and R_2 have the same significance as in Fig. 11-10. Actually μ and r_p (especially the latter) are not constant, but

[1] See R. L. Freeman, Use of Feedback to Compensate for Vacuum-tube Input-capacitance Variations with Grid Bias, *Proc. IRE*, 26, p. 1360, November, 1938; and John F. Farrington, Compensating for Tube Input Capacitance Variation by Double Bias Provision, *Communications*, 20, p. 3, September, 1940.

quite satisfactory conclusions may be drawn concerning the frequency stability of the oscillator by assuming average values for these tube coefficients.

Let the current $\mathbf{I}_2$ flowing through coil L_2 be the reference vector (Fig 11-11*b*). The voltage across this coil will lead the current by slightly less than 90 deg owing to losses in the circuit, *i.e.*, it will be equal to $\mathbf{I}_2(R_2 + jX_2)$ and is shown as $\mathbf{E}_p$, Fig 11-11*b*.

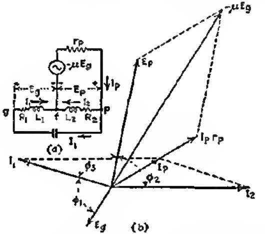


FIG 11-11. (*a*) Equivalent a-c circuit and (*b*) vector diagram of a Hartley oscillator ($1/\omega C_g = 0, \omega L_g = \infty$). ϕ_1, ϕ_2, and ϕ_3 are all slightly less than 90 deg

The current flowing through the tank condenser C must lead the voltage $\mathbf{E}_p$ by slightly less than 90 deg; *i.e.*,

$$\mathbf{I}_1 = \frac{\mathbf{E}_p}{R_1 - j(X_c - X_1)} \tag{11-38}$$

The reactance X_1 is of course smaller than the capacitive reactance X_c, so the net reactance is capacitive. The voltage $\mathbf{E}_g$ across the terminals of coil L_1 will lead the current $\mathbf{I}_1$ by a little less than 90 deg, being equal to $\mathbf{I}_1(R_1 + jX_1)$. This same voltage will be applied to the grid of the tube, since the drop through the grid condenser was assumed to be negligibly small. It may therefore be seen that the grid and plate voltages are *not out of phase by* 180 *deg* but are shifted through a small angle *owing to the resistance*

components of the tank circuit. The only possible way for them to differ by 180 deg is for R_1 and R_2 to equal zero which would mean zero output as well as zero losses.

Inspection of Fig. 11-11*b* shows that any change in the length of the $\mathbf{I}_p r_p$ vector (due to a change in r_p) must be accompanied by a change in the relative positions of the $\mathbf{E}_p$ and $\mathbf{E}_g$ vectors to satisfy the relation

$$-\mu \mathbf{E}_g = \mathbf{I}_p r_p + \mathbf{E}_p \tag{11-39}$$

This phase shift is provided automatically by a slight change in frequency sufficient to alter the reactances of the tank circuit by the requisite amount.

Methods of Increasing Frequency Stability. An increase in Q always improves frequency stability. This may be seen by recalling that high values of Q indicate large energy storage, and large energy storage in a tank circuit has a similar effect to the storage in the flywheel of a reciprocating engine which tends to maintain constant angular velocity of the wheel and therefore sinusoidal motion of the crosshead. The effect of Q is also indicated by Eq. (10-42) where an increase in Q is seen to be accompanied by a decrease in the tank circuit reactance. The resistances and capacitances of the tube are seen to be in parallel with the tank circuit reactance and will necessarily have less effect as this reactance is decreased.

Frequency stability may be improved by the insertion of suitable reactances in series with the plate or grid leads to bring $-\mu \mathbf{E}_g$ and $\mathbf{E}_p$ into phase, since a change in r_p will then require a change of $\mathbf{E}_p$ and $\mathbf{E}_g$ in amplitude only, not in phase. In the Hartley circuit the plate-blocking condenser will serve this function if properly designed. For circuits with a Q of approximately 12, C_p should be of the same order of magnitude as the tank capacitance, for most triodes.[1]

Since changes in load, either resistive or reactive, affect the frequency, oscillators should be operated into a constant load. One method of improving the constancy of the load is to use the oscillator to drive a class C amplifier with the load coupled to the plate

[1] For more details see D. C. Prince and F. B. Vogdes, Vacuum Tubes as Oscillation Generators, Part IV, *Gen. Elec. Rev.*, **30**, pp. 147–152, March, 1928; also F. B. Llewellyn, Constant Frequency Oscillators, *Proc. IRE*, **19**, p. 2063, December, 1931.

circuit of this amplifier. Such a combination is called a *master oscillator, power amplifier* (abbreviated MOPA). For maximum stability the grid of the class C amplifier should not be driven positive and, when so operated, the amplifier is commonly known as a *buffer amplifier*. The buffer normally is of relatively low-power output and is generally used to drive a class C amplifier of normal design. This second amplifier supplies the load or drives another amplifier of still higher power rating.

Multigrid tubes may serve as both oscillator and buffer, as in the circuit of Fig. 11-12, known as an *electron-coupled oscillator*. Here the first two grids of a pentode constitute the control grid and anode of a conventional oscillator, and the plate supplies the load. The electron stream flowing through the tube to the plate is varied by the oscillating potentials on the first two grids, pro-

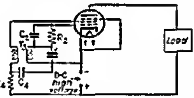


FIG. 11-12 Circuit of an electron coupled oscillator using a pentode tube

ducing a current in the load of the same frequency. The third grid acts as a screen practically to eliminate any capacitance between the plate and the first two grids of the tube. Thus a change in load will have no effect on the frequency of oscillation, since the electron stream that supplies the output is unidirectional and all tube capacitances between the plate and the first two grids have been virtually eliminated. It is evident, of course, that no external capacitances or other sources of coupling between the load and the oscillatory circuits can be tolerated if maximum frequency *stability is to be realized.*

The screen grid of the tube in the circuit of Fig. 11-12 provides no screening action since its potential varies with the frequency of oscillation, all shielding of the plate being due to the suppressor grid alone. In the circuit of Fig. 11-13 the screen grid, while serving as an anode for the oscillator portion of the circuit, is maintained at ground potential for radio frequency and so pro-

vides screening action between the plate and the oscillator portion of the tube. Since only one electrode may be grounded, the cathode must be operated above ground for radio frequency as shown. The capacitance between cathode and heater is effectively across points K and P but is sufficiently small to have negligible effect at any but the higher radio frequencies. Its effect may be removed by inserting r-f chokes in each heater wire at the points indicated by crosses.

Grid-condenser Design. If the grid condenser is not called upon to correct for phase angle between grid and plate voltage, its exact size is not critical. It must be sufficiently large to maintain the grid bias reasonably constant throughout the cycle; *i.e.*, the time constant of the grid condenser and grid leak must be fairly long as compared to the time of one cycle. However, it should not be made any larger than necessary, as too large a grid

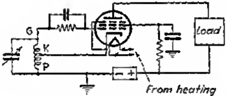


FIG. 11·13. Electron-coupled oscillator with improved shielding of the plate.

condenser may cause the oscillator to operate intermittently, the intervals between oscillating periods being a function of the time constant of the circuit. This may be explained as follows:[1]

When the tube is operating at maximum efficiency, it is necessarily operating fairly close to the point of instability, so that a momentary drop in alternating grid voltage, due to a sudden increase in load or other disturbance, may cause the output from the tube to become less than the power absorbed by the load. The oscillations in the tank circuit will then tend to die down, further decreasing the alternating grid voltage. With a small-sized grid condenser, however, the grid bias will immediately decrease, producing an increased plate-current pulse, and thereby bring opera-

[1] For a more complete discussion of this phenomenon see Prince and Vogdes, *loc. cit.*

tion back to normal. But if the grid condenser is too large, it will tend to maintain the grid-bias constant so that any drop in alternating grid voltage will decrease the size of the plate-current pulse, thereby further lowering the power supplied to the tank circuit until the oscillations cease. After a short interval of time the charge will leak off the grid condenser, and the tube will again start oscillating, building up to its normal output, after which the cycle may repeat. For oscillators with a Q of the order of 12 to 20, a grid condenser having a capacitance of about 50 per cent that of the tank condenser should prove satisfactory. If intermittent oscillations do occur, a decrease of either the grid condenser or the grid leak should correct the condition.

R-C Oscillator. Although most oscillators are built with both inductance and capacitance to provide energy storage and to determine the frequency of oscillation, it is possible to build one with capacitance and resistance (no inductance), known as an *R-C oscillator*. Such an oscillator has the advantages of greater simplicity in construction (inductances must normally be shielded) and of greater frequency stability and generally better wave shape. The frequency range is limited to that in which a resistance-coupled amplifier will operate; thus R-C oscillators operate at audio or at video frequencies, not at higher radio frequencies.

Figure 11-14 shows a block diagram of an oscillator of this type. It consists of a resistance-coupled amplifier with negative feedback but with an added *positive* feedback circuit consisting of $R_1R_2C_1C_2$. The positive feedback varies with frequency somewhat as indicated by curve a of Fig. 11-15 while the negative feedback is relatively independent of frequency, as in curve b of Fig. 11-15. Thus if the negative feedback is less than the maximum positive feedback, as indicated, the amplifier will oscillate at the frequency of maximum positive feedback. If the negative feedback is made appreciably less than that which will just permit the amplifier to oscillate, the wave shape will be poor; if made large enough so that the amplifier is just able to oscillate, the wave shape is remarkably good with a distortion as low as a fraction of 1 per cent. In a commercial unit special means are provided to ensure that the proper amount of negative feedback is maintained at all frequencies (as in the circuit of Fig. 11-16, for example, which is described on page 466).

The circuit of Fig. 11-9 may be used to represent this oscillator

even though no actual inductance is used. To demonstrate this let us first determine the impedance seen looking to the right at the terminals FG in Fig. 11-14. This may be done by dividing the voltage $\mathbf{E}$ by the current flowing at F. Since the input impedance of the resistance-coupled amplifier is normally very high, this current may be considered as being equal to the current $\mathbf{I}$ flowing through the positive feed-back circuit. Adding the potentials around this circuit from ground on the input to ground again at the output, we may write

$$\mathbf{E} - \mathbf{I}\left(R_1 + \frac{1}{j\omega C_1} - \mathbf{AE}\right) = 0 \qquad (11\text{-}40)$$

where $\mathbf{AE}$ is the output voltage, using the same notation as was

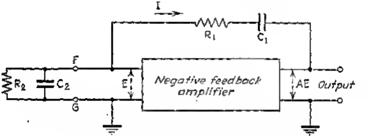


FIG. 11-14. Block diagram of a R-C oscillator.

used in developing the equations for the resistance-coupled amplifier, page 269.

Equation (11-40) may now be solved for the voltage $\mathbf{E}$, and the impedance may then be found by dividing $\mathbf{E}$ by $\mathbf{I}$ to give

$$Z_{FG} = \frac{\mathbf{E}}{\mathbf{I}} = \frac{1 + j\omega R_1 C_1}{j\omega C_1(1 - \mathbf{A})} \qquad (11\text{-}41)$$

or

$$Z_{FG} = \frac{R_1}{1 - \mathbf{A}} - j\,\frac{1}{\omega C_1(1 - \mathbf{A})} \qquad (11\text{-}42)$$

A comparison of Figs. 11-14 and 11-9 will show that R_2 and C_2 of the former figure could represent r and C of the latter figure, in which case Z_{FG} should represent R and L of Fig. 11-9. For this to be so, the reactive term of Z_{FG} must be inductive instead

of capacitive which will be the case if **A** has a zero phase angle and is greater than unity. Furthermore Eq. (11-14) shows that either r or R must be negative if α is to be zero. If **A** has a zero phase angle and is greater than unity, the resistive component of Z_{ro} will be negative, thus fulfilling all the requirements for oscillation as set up for Fig. 11-9.

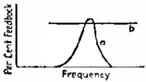


FIG. 11-15 Illustrating the manner in which feedback varies with frequency in a R-C oscillator

Reference to Eq. (9-16), page 269, shows that **A** for a single-stage, resistance-coupled amplifier is, throughout the useful frequency range of the amplifier, greater than unity but has a phase angle of 180 deg. It is therefore necessary to use two stages of amplification in order that the over-all gain factor shall have a zero phase angle, and **A**, as used here, represents this over-all factor and not the gain per stage as in Chap. 9

Since **A** is to have zero phase angle and be greater than unity, it seems desirable to rewrite Eq. (11-42) in terms of $(\mathbf{A} - 1)$ instead of $(1 - \mathbf{A})$ and to use the magnitude A instead of the vector **A**. This gives

$$Z_{ro} = -\frac{R_1}{A-1} + j\frac{1}{\omega C_1(A-1)} \tag{11-43}$$

This form of the equation shows more clearly than Eq (11-42) that the amplifier and positive feed-back circuit R_1C_1 are the equivalent of a negative resistance R and an inductance L, where

$$Z_{ro} = R + j\omega L \tag{11-44}$$

$$R = -\frac{R_1}{A-1} \tag{11-45}$$

$$L = \frac{1}{\omega^2 C_1(A-1)} \tag{11-46}$$

It is evident that the equivalent inductance L of Eq. (11-46) is quite different from the usual inductance since it varies inversely as the square of the frequency. This is because its reactance varies *inversely* as the frequency, not *directly*, as indicated by the second term of Eq. (11-43)

One might expect that the frequency of oscillation would be

found by substituting into Eq. (11-29), but instead the frequency is found from Eq. (11-14) by making α equal to zero. Making the substitution of $r = R_2$, $C = C_2$ and substituting for R and L from Eqs. (11-45) and (11-46) gives

$$0 = \frac{R_1}{A-1}\frac{\omega^2 C_1(A-1)}{2} - \frac{1}{2R_2C_2} \tag{11-47}$$

Solving for ω gives

$$\omega = \frac{1}{\sqrt{R_1R_2C_1C_2}} \tag{11-48}^*$$

If, as is normally the case, $C_1 = C_2 = C$ and $R_1 = R_2 = R$, Eq. (11-48) may be rewritten in terms of frequency as

$$f = \frac{1}{2\pi RC} \tag{11-49}$$

The capacitances C_1 and C_2 are normally variable, air-dielectric condensers mounted on the same shaft, and the frequency is varied by turning the shaft of this dual condenser. A comparison of Eq. (11-49) with Eq. (11-34) for tuned-circuit oscillators shows an advantage of R-C oscillators to be that, as the condenser is rotated, the frequency varies inversely as C, not as the square root of C, which gives a greater frequency change from minimum to maximum condenser positions. In addition, the range of frequencies covered may be changed by adjusting R in steps so that the R-C oscillator may easily be designed to cover any band of frequencies for which a resistance-coupled amplifier may be satisfactorily designed.

It is of interest to substitute R_1, R_2, C_1, and C_2 into Eq. (11-29). If the equation is first squared, the result is

$$\omega^2 = \frac{\omega^2 C_1(A-1)}{C_2}\,\frac{R_2 - R_1/(A-1)}{R_2} \tag{11-50}$$

We may cancel ω out of both sides of the equation and simplify to obtain

$$A = \frac{C_2}{C_1} + \frac{R_1}{R_2} + 1 \tag{11-51}$$

This is the condition required for steady-state operation and, if A tends to exceed this value, the amplitude of oscillations will in-

crease until operation takes place over the curved portions of the static characteristic curves of the tubes. This will increase the plate resistance of the tubes until A is decreased sufficiently to satisfy the requirements of this oscillator but will produce nonsinusoidal output. To ensure good wave shape, the negative feedback of the amplifier must be adjusted until the amplifier will just oscillate whence A will very nearly satisfy Eq. (11-51) with the positive feedback removed. For best performance this action should be automatic.

If $R_1 = R_2$ and $C_1 = C_2$, as is usual, A must evidently be equal to 3.

Figure 11-16 shows the circuit of a commercial type of R-C

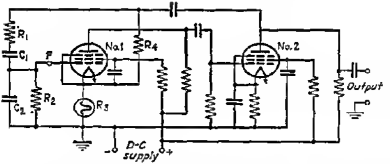


FIG. 11-16 Circuit of a R-C oscillator. R_3 is a small tungsten lamp.

oscillator. If R_1, R_2, C_1, and C_2 were removed, the remainder of the circuit would be a conventional two-stage, resistance-coupled amplifier with negative feedback, the input being between terminal F and ground. The negative feed-back circuit is through R_4 and R_3, this being essentially the same circuit as that of Fig. 9-74, page 370. The resistance R_3 is commonly a small tungsten lamp, the resistance of which varies widely with the magnitude of the current flowing through it. Thus if the amplitude of the oscillations starts to increase, R_3 becomes larger, increasing the negative feedback and preventing any appreciable rise in amplitude. This procedure automatically maintains the negative feedback at approximately its maximum permissible value as R_1, R_2, C_1, and C_2 are changed to adjust the frequency, and so maintains reasonably good wave shape at all frequencies.

Phase-shift Oscillator. Figure 11-17 shows the circuit of a single-tube, resistance-capacitance oscillator known as a *phase-shift oscillator*. Appreciable phase shift is introduced by each resistance-capacitance mesh (*RC* in the figure) so that, although the tube introduces a phase shift of 180 deg, the over-all gain A from *F* to *H* can have a phase angle of zero with an amplitude equal to or greater than one. The amplifier will then tend to oscillate at the frequency for which the phase angle of A is zero and therefore at the frequency for which the phase shift through the resistance-capacitance network is 180 deg. Good wave shape requires adjustment of the amplifier gain until oscillations are barely maintained; *i.e.*, A should equal one.

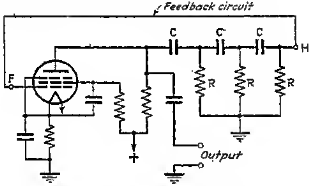


FIG. 11-17. Phase-shift oscillator.

Crystal Oscillators. In radio transmitters and other applications, the frequency of the output signal must be maintained constant to a very high degree. In radio broadcasting, for example, stations are assigned to certain channels, and the various channels are spaced 10 kc apart. Since the frequency spectrum covered by a broadcast station is not less than 10,000 cycles/sec (owing to the side bands, see Chap. 13), it is evident that the transmitted frequency of a broadcast station must be held constant to a very close tolerance, the present figure being 20 cycles/sec.[1] The

[1] Another factor affecting the frequency tolerance required of broadcast stations is the production of beat frequencies by stations located some distance apart but assigned to the same frequency. The physical spacing of the stations is normally such as to preclude the possibility of any appre-

frequency tolerance for stations in other types of service is not quite so rigid but must not exceed 0.05 per cent for most services. To maintain this close tolerance, piezoelectric crystals are commonly used in the tank circuits of the frequency-controlling oscillators.

Certain crystalline materials, notably Rochelle salts, quartz, and tourmaline, are capable of producing piezoelectric effects, *i.e.*, when compressed or otherwise placed under mechanical strain, electric charges appear on opposite faces of the crystal. Conversely, when electric charges are applied to opposite faces, mechanical strains are set up within the crystal. The magnitude of the potential set up by mechanical strain varies from a small fraction of a volt up to as high as several hundred volts in Rochelle salts.

This piezoelectric effect may be utilized to fix the frequency of a vacuum-tube oscillator with an accuracy of better than 1 part in 1,000,000. A crystal is so ground that, when set in vibration by suitable application of alternating potentials, mechanical resonance will occur at the desired frequency. Quartz crystals are generally used for this purpose, although they produce less response than Rochelle salts, they are much better mechanically, being less liable to fracture under strong electrical fields or mechanical shock. Tourmaline is still better mechanically but weaker electrically than quartz and finds some application at very high frequencies, where the crystals must be ground so thin as to make the use of quartz impractical.

A simple circuit suitable for use with a piezoelectric crystal is shown in Fig. 11-18. The circuit is essentially that of the tuned-grid, tuned-plate oscillator of Fig. 11-7, the crystal and its holder constituting the tuned-grid circuit. This circuit will oscillate when sufficient energy is fed back to the grid circuit to supply its losses. Referring to Eq. (9-7), the power flowing *from the grid circuit* through the grid-plate capacitance is

---

ciable reception of the modulation of both stations, but there may be regions between the stations where the two carriers are of such strength as to produce a beat note or whistle the frequency of which is equal to the difference in the frequencies of the two stations. Since the frequency tolerance is but 20 cycles/sec, the maximum beat frequency (40 cycles/sec) is too low

[illegible] , and it is extremely [illegible] extremes of their per-

$$P = E_g^2 G = E_g^2(-A\omega C_{gp} \sin \psi) \qquad (11\text{-}52)$$

where G is that component of the input conductance due to the current flowing through the grid-plate capacitance. If the right-hand term of this equation is negative, the power flow must be in the opposite direction, or *toward the grid*, a necessary condition for oscillation. The term will be negative if $\sin \psi$ is positive (inductive load) and will be large if the load impedance is high.[1]

The high inductive reactance required in the plate circuit of the crystal oscillator may be supplied by a large inductance but can generally be more cheaply and satisfactorily obtained by using a parallel condenser and inductance tuned to slightly above the resonant frequency of the crystal. The condenser in the plate circuit of Fig. 11-18 is therefore adjusted until the crystal oscillates satisfactorily; it has but little effect on the frequency of oscillation.

The output of crystal-controlled oscillators is limited by the alternating voltage that the crystal will withstand. Excessive voltage may cause the crystal to vibrate mechanically with such violence as to fracture. The low output from this type of oscillator is no limitation on its usefulness, since it may be used to drive suitably neutralized or screen-grid amplifiers, amplifying the signal to any desired level. In the radio field, crystal oscillators are used to control transmitters delivering hundreds of kilowatts into the antenna.

FIG. 11-18. Simple circuit of a piezo-electric crystal oscillator.

Characteristics of Quartz Crystals. Natural quartz crystals are hexagonal in shape with pointed ends (Fig. 11-19), although they are seldom found with both ends perfect. Three principal axes have been designated, the Z axis being known as the optical axis and the X and Y axes as, respectively, the electrical and mechanical axes. The X axis, as may be seen from the figure, runs between opposite points of the hexagonal cross section of the crystal, whereas the Y axis is perpendicular to the opposite faces. Crystals cut perpendicular to the X axis are known as X-*cut* crystals; those cut perpendicular to the Y axis are known as Y-*cut*, or 30-*deg cut*.

[1] This follows since A increases with the load impedance.

Crystals may be made to vibrate at a frequency that is a function of their width or of their thickness, as desired. The thickness of an X-cut crystal is in the direction of an X axis, and its width is in the direction of a Y axis.[1] Similar relations hold for the Y-cut crystal except that its width is along the X axis and its thickness in the direction of the Y axis. For very low frequencies, say below 200 kc, the width vibration is used, as the mechanical dimensions for such low frequencies are comparatively large, about 2.86 cm being required for a frequency of 100 kc. Width-vibration crystals are usually so cut that their longest dimension is along the Y axis (for an X-cut crystal), this dimension determining their frequency of oscillation. For higher frequencies, thickness vibrations are generally employed in order that the dimensions of the

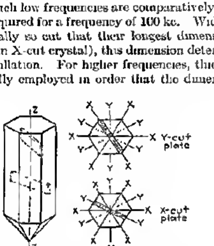


FIG. 11-19. Sketch of quartz crystal showing location of X and Y cuts. Actual crystals are not so regular as the one shown.

crystal may not be too small. For this type the surfaces must be accurately ground to secure the correct thickness at every point in the crystal.

Zero-temperature-coefficient Crystals.[2] The resonant frequency of both X- and Y-cut crystals varies with the temperature. The temperature coefficient of all X-cut crystals and of width-vibration Y-cut crystals is negative, whereas that of thickness vibration Y-cut crystals may be either positive or negative or even zero de-

[1] This may be considered as its length if the dimension along the Y axis exceeds that along the Z axis. This is commonly the case for crystals operating at very low frequencies which are often referred to as *crystal bars*.

[2] An excellent discussion of such crystals is given by W. P. Mason, Low Temperature Coefficient Quartz Crystals, *Bell System Tech. J.*, **19**, p. 74, January, 1940.

pending on the temperature and the relative dimensions of the crystal. Y-cut crystals are not used extensively, as they often have more than one frequency of oscillation, differing from each other by only a small percentage. These multiple frequencies are especially prevalent in thin crystals. X-cut crystals have been widely used in temperature-controlled ovens which maintain the temperature of the crystal to within 0.1°C, usually at a temperature of about 50°C. When so operated, they are capable of maintaining the frequency constant with an accuracy of as high as 1 part in 10,000,000.

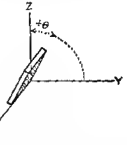


FIG. 11-20. Illustrating the method of cutting a crystal at an angle to the Z axis to secure a zero-temperature coefficient.

More recently crystals have been cut at an angle to the Z axis by rotating the cutting plane around the X axis (Fig. 11-20). It has been found that a rotation of about 35 deg in a positive direction (clockwise in Fig. 11-20) will produce a crystal with zero-temperature-coefficient, known as an *AT cut* (Fig. 11-21). If the plane is rotated 49 deg in the opposite direction, another zero-temperature-coefficient crystal is obtained known as a *BT-cut* crystal. These crystals vibrate in shear, similar to the Y-cut crystals.

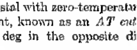


FIG. 11-21. AT and DT cuts. θ should be about 35 deg.

Two other zero-temperature-coefficient crystals are obtained from what are known as the CT and the DT cuts. The CT cut is approximately at right angles to the BT cut, and the DT approximately at right angles to the AT (Fig. 11-21). These are especially suited for lower frequency performance, notably in the range of 50 to 500 kc. They oscillate along their longer dimensions in such a way that one pair of opposite corners draws in toward the center of the crystal while the other pair moves outward. This is illustrated in Fig. 11-22 where the dotted lines

indicate, in a greatly exaggerated manner, the movement of the crystal during one-half cycle of oscillation.

All the zero-temperature-coefficient crystals just described exhibit their constant-frequency properties at one temperature only, the temperature coefficient increasing from zero as the temperature is varied. Thus for maximum accuracy some sort of temperature-controlling oven is needed even with these crystals, although the error introduced by a small change in temperature is very much less than in X-cut crystals, and an improvement in frequency stability may be expected either with or without an oven.

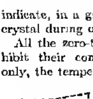


Fig. 11-22. Illustrating the mode of oscillation of CT and DT crystals

Still another type of cut produces a crystal in which the temperature coefficient is essentially zero over a very wide range of temperature, the frequency varying less than 1 part in 1,000,000 throughout a temperature range of 100°C This crystal is obtained by rotating the principal axes of a CT or DT crystal by 45 deg, as shown in Fig. 11-23, and is known as a *GT crystal*[1] By comparison with Fig 11-22 the vibration will be seen to consist of longitudinal vibrations as shown by the dotted lines in Fig 11-23.

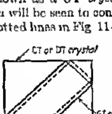


Fig. 11-23 Showing how the GT cut is taken from a CT or a DT cut.

The vibration system of any crystal is extremely complex. Each crystal, for example, has at least two vibrating frequencies: one a function of its width and the other of its thickness Coupling exists in the crystal between these different modes of vibration and between their harmonics. The degree of coupling varies with the dimensions of the crystal and with the relative values of the different vibration frequencies, often producing multiple resonant frequencies at closely spaced intervals. This is particularly true of Y-cut crystals and is one of the reasons why this cut has not been widely used In the special zero-temperature-coefficient crystals one of the two principal modes of vibration has a positive temperature coefficient and the other has a negative coefficient.

[1] W. P. Mason, A New Quartz-crystal Plate, Designated the GT, Which Produces a Very Constant Frequency over a Wide Temperature Range, *Proc IRE*, 28, p 220, May, 1940

interaction between the two gives a resulting vibration period independent of temperature.

The ratio of the axes in the GT crystal determines the degree of coupling between the two principal modes and therefore controls the temperature coefficient. This ratio may be varied by grinding one edge of the crystal until the coefficient is zero. It has been found that a zero coefficient may be obtained in this manner at angles of cut lying between +35 and +75 deg. For negative angles the range lies between −50 deg and an angle somewhat greater than −70 deg. Mason reports use of a GT crystal for measurements in the Caribbean Sea with no temperature control but with the crystal contained in an evacuated container to exclude moisture and the effects of barometric pressure, wherein the frequency was maintained constant to within several parts in 10,000,000.[1]

The GT is used at low frequencies, its maximum frequency being about 1000 kc. For higher frequencies the AT cut is most commonly employed, although, as previously stated, maximum precision requires some sort of temperature control. Nevertheless, the frequency stability without temperature control is sufficiently high for many applications, and the crystal is often so used.

Equivalent Circuit of the Crystal. Crystals are used in vacuum tube oscillators by mounting them between two electrodes connected to suitable points in the oscillator circuit. Since the mechanical vibrations of the crystal induce electric charges on these electrodes, the crystal may be replaced by an equivalent electrical resonant circuit having suitable constants.[2] Figure 11-24 shows such an equivalent circuit in which C_m represents the capacitance between the electrodes and any capacitance in the external circuit; and L, R, and C represent, respectively, the mass, the mechanical and molecular friction, and the compliance of the crystal. These constants have been determined for various crystals and show the crystal to be the equivalent of an electrical circuit having a Q higher than can possibly be obtained with any coil and condenser

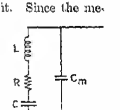


FIG. 11-24. Equivalent circuit of a crystal and its holder.

[1] Mason, *loc. cit.*

[2] K. S. Van Dyke, The Piezoelectric Resonator and Its Equivalent Network, *Proc. IRE*, 16, p. 742, June, 1928.

combination yet designed. A typical set of constants for a 90-kc crystal is L, 135 henrys; C, 0.021 $\mu\mu$f, R, about 7500 ohms. The Q of such a crystal is therefore over 11,000, enormously higher than any high-grade electrical circuit. Some of the more recently developed GT crystals have a Q as high as 330,000 when properly mounted in an evacuated container. Higher frequency crystals generally have a somewhat lower Q, but one that is still sufficiently high to ensure precision of frequency control. As explained on page 459, Q is the figure of merit in determining the constancy of frequency.

The constancy of the crystal-oscillator frequency may be illustrated in another way. The crystal itself is the equivalent of a series-resonant circuit (Fig. 11-24), whereas the tuned-grid, tuned-plate oscillator (of which the circuit of Fig. 11-18 is the equivalent) requires a parallel-resonant circuit between grid and cathode. In the circuit of Fig. 11-24 the parallel-resonant circuit consists of C_m in parallel with an equivalent inductance from the crystal, *i.e.*, the crystal operates at a frequency slightly above its resonant point so that its net reactance is inductive. Assume that a crystal having the constants given in the preceding paragraph is mounted in a holder with a capacitance C_m equal to 3.5 $\mu\mu$f. The crystal will oscillate at the frequency at which the reactance of the crystal circuit is equal to that of C_m, or very nearly so:

$$\omega \times 135 - \frac{10^{12}}{\omega \times 0.024} = \frac{10^{12}}{\omega \times 3.5}$$

Solving for ω gives

$$\omega = \sqrt{\frac{0.286 \times 10^{12} + 41.6 \times 10^{12}}{135}} = 557 \times 10^3$$

567015.85

or

$$f = 88.7 \text{ kc}$$

The important point in the foregoing operation is that the first term under the radical represents the effect of the external capacitance C_m, which may very evidently be increased or decreased considerably with almost negligible effect on the frequency. The reactance of C_m at resonance is about 0.5 megohms, whereas that of L and C for the crystal are 76.3 and 75.8 megohms, respectively. Consequently an increase of 100 per cent in C_m would require only sufficient shift in frequency to make X_L and X_c for the crystal

differ by 1 megohm instead of 0.5, or $X_L = 76.55$ and $X_c = 75.55$. This represents a frequency shift of 0.30 per cent.

This change in C_m may be purposely made to produce a slight change in frequency, or it may be an equivalent change due to reaction from the plate circuit as represented by the last term of Eq. (9-7) (page 262). In either event the crystal prevents any very appreciable change in frequency.

Bridge-stabilized Oscillator. Crystals are capable of giving increased frequency stability when used in a suitable bridge with a class A amplifier of high gain. Such a unit is known as a *bridge-stabilized oscillator*, and its circuit is shown in Fig. 11-25.[1] Three of the four arms of the bridge are purely resistive, and the fourth is a crystal. The crystal operates at its series-resonant frequency (see the equivalent circuit of Fig. 11-24) so the shunt capacitance C_m of its mounting has a nearly negligible effect. The amplifier is a conventional class A unit designed to produce negligible distortion.

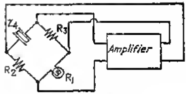


FIG. 11-25. Circuit of the bridge-stabilized oscillator.

Evidently when the bridge is balanced, no signal is fed back to the input of the amplifier and the circuit is inoperative. Thus the bridge must run slightly off balance, the amount of unbalance being such that the loss inserted by the bridge just offsets the gain of the amplifier. In practice one of the arms R_1 consists of a resistance with a very high positive temperature coefficient, such as a small tungsten lamp, so that unbalancing of the bridge is automatically taken care of. If the attenuation of the bridge is too small, the output of the amplifier increases, which in turn increases the resistance of the arm R_1 until balance is restored. Too high a resistance at R_1 will produce the opposite reaction. Since the variation in resistance of this arm is comparatively slow, it will not follow the r-f currents impressed on the bridge, and thus no nonlinearity is introduced to cause the generation of harmonics or their intermodulation.[2]

[1] See L. A. Meacham, The Bridge-stabilized Oscillator, *Proc. IRE*, 26, p. 1278, October, 1938.

[2] Llewellyn has shown that this is one of the causes of frequency instability; see footnote on p. 459.

Any change in phase angle or gain in the amplifier will cause a small shift in frequency, but the bridge circuit limits this shift to a very small value. Furthermore, the frequency stability increases with the gain of the amplifier. Consequently an oscillator of extremely high frequency stability may be built by using a high-gain amplifier and a low-temperature-coefficient crystal operated in a constant-temperature oven. Two such oscillators were constructed and operated over a period of several months with a maximum frequency variation of about 2 parts in 100,000,000, except for a slow drift due to aging of the crystals. This drift may be corrected by inserting an adjustable reactance in series with the crystal. Slight variations in this reactance will not appreciably affect the stability.

Uses of Crystal Oscillators. Probably the largest single application of crystal oscillators is in maintaining the frequency of radio transmitters within the tolerance requirement of the Federal Communications Commission. Other applications are found in calibration work. A crystal oscillator, for example, may be used to drive a multivibrator (to be described in a succeeding section of this chapter) by means of which frequencies having any desired multiple or submultiple of the oscillator may be obtained. If operated in conjunction with an accurately calibrated audio oscillator, any radio frequency whatsoever may be obtained, with an accuracy dependent only upon the quality of the crystal and the degree of its temperature control.[1]

In addition to these two fields there are many special applications where the high constancy of frequency desired may be achieved only through the use of a crystal oscillator.

Frequency Control with Resonant Lines.[2] The frequency of an oscillator may be held constant by the use of a resonant line in place of the conventional tank circuit. For example, a transmission line that is an odd number of quarter wave lengths long and short-circuited at the far end acts like a parallel-resonant circuit with a Q much higher than that of the ordinary coil and

[1] See Frederick E. Terman, "Measurements in Radio Engineering," McGraw-Hill Book Company, Inc., New York, 1935.

[2] An excellent discussion of this method of frequency control at only moderately high frequencies is given by C. W. Hansell and P. S. Carter, Frequency Control by Low Power Factor Line Circuits, *Proc. IRE*, 24, p. 597, April, 1936.

condenser type of tank circuit. Such a tank circuit is, of course, most adaptable to u-h-f oscillators, because the physical dimensions of the line become cumbersome as the frequency is decreased. At a frequency of 300 Mc, for example, a quarter wave length is only a little over 0.8 ft long, and the dimensions are correspondingly smaller at higher frequencies.

Lighthouse tubes and their associated circuits present an excellent example of the use of transmission lines as resonant circuit elements for r-f oscillators. Figure 11-26 shows the lighthouse tube of Fig. 3-62 with the necessary transmission lines to cause it to perform as an oscillator. The oscillator consists of a resonant circuit between plate and grid, formed by the concentric transmission line between sleeves *A* and *B*, and a second resonant circuit between grid and cathode consisting of the transmission line formed by sleeves *B* and *C*. Feedback must be provided between the cavities of these two lines, as by the coupling loop indicated in the figure, to cause oscillations. The frequency is determined by the length of the sleeves, and the tube may be operated at frequencies up to about 1500 Mc. Direct potentials are supplied to the electrodes by connecting external leads to the ends of the sleeves through suitable chokes.

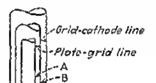


FIG. 11-26. Lighthouse tube with concentric transmission lines for resonant circuits.

Dynatron. A negative-resistance characteristic can be used to produce oscillations by what is known as *dynatron* action. The significance of a *negative* resistance is that it represents the flow of power *into* a circuit instead of *out*. Referring to Figs. 3-40 and 3-44 (pages 73 and 76), the plate resistance of a screen-grid tetrode tube is seen to be negative when the plate is maintained at a potential somewhat lower than that of the screen. Consequently the screen-grid tube can be used to provide a flow of power into a suitable circuit if the electrode voltages are adjusted to cause operation in this negative region. Figure 11-27 shows a simple circuit utilizing this principle in which the plate is supplied through a resonant circuit and is maintained at a potential somewhat lower than that of the screen grid. The function of the

resonant circuit is, of course, to provide energy storage for maintaining oscillations, just as in the conventional type of oscillator.

The detailed operation of the dynatron may be understood from the curves of Fig. 3-40, one of which is reproduced in Fig. 11-28. The electrode voltages should be adjusted to cause the

FIG. 11-27 Basic circuit of a tetrode used as a dynatron oscillator

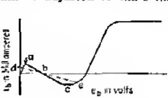


FIG. 11-28 Static i_b-e_b curve of a tetrode reproduced from Fig 3-40

normal, static plate current to lie near the middle of the negative region, such as at b, Fig. 11-28. A momentary rise in plate current will evidently cause a drop in plate voltage, and since the battery voltage is constant, any drop in plate voltage must be accompanied by a rise in the voltage across the tank circuit, causing the plate current to rise still farther beyond its normal value b. This process will continue until the current approaches its maximum value at a, when it will cease increasing, causing the voltage across the tank circuit to decrease again. The tube drop will then rise, and the plate current will fall to its minimum value c, the polarity of the voltage across the tank circuit being opposite to what it was during the building-up process. This cycle will be repeated at a frequency determined by the tank inductance and capacitance.

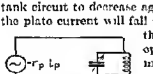


FIG. 11-29 Equivalent circuit of a dynatron oscillator.

The equivalent circuit of the dynatron is shown in Fig. 11-29 where the alternator represents the negative resistance of the plate circuit of the tube. Its voltage will be $e_p = -r_p i_p$, where r_p must be negative. The circuit will oscillate if the minimum negative plate resistance, as represented by the slope of the curve at point b, is less than the resistance presented by the tank circuit; but the final amplitude of the oscillations must be such that the *average* value of the negative plate resistance becomes equal to the tank resistance. A high-impedance tank circuit will, therefore,

cause the tube to oscillate well over the bends of the curve, as to points d and e (Fig. 11-28), where the dotted line gives some indication of the resulting increase in average r_p.

Good frequency stability and wave shape may be secured with dynatron oscillators by limiting their operation to the region abc, Fig. 11-28, since their performance will then be similar to that of class A oscillators. A simple method of securing the desired range of operation is to vary the control-grid bias, as this will alter the slope of the static curve (see Fig. 3-40, page 73) and thus vary r_p. Maximum stability requires that the grid bias be increased to the most negative potential at which oscillations are sustained. The excellent frequency stability and good wave shape thus obtainable, together with the simplicity of the circuit, make this type of oscillator of considerable value in the laboratory.

Negative Transconductance Oscillator. A principal disadvantage of the dynatron oscillator is its dependence upon secondary emission, which varies considerably with age in any given tube. Similar oscillations are obtainable without the need of secondary emission by operating a pentode in such a manner that the transconductance between two of its grids becomes negative. A tube operating in this manner has been termed a *transitron* oscillator.[1]

The principle of operation of this oscillator is illustrated in Fig. 11-30. The potential E_0 is sufficient to make the grid G_3 more negative than the cathode. Grid 3 therefore repels electrons passing through grid 2 (here used as an anode) and forms a virtual cathode (see discussion of beam tubes, page 83, for definition of a virtual cathode). During the negative half cycle of the voltage across the tank circuit the potential of the anode grid G_2 is decreased, and less current might be expected to flow. However, the tank voltage also

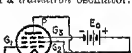


FIG. 11-30. Circuit illustrating the principle of operation of the negative transconductance oscillator.

increases the negative voltage applied to grid G_3, causing it to repel electrons more effectively and actually increasing the current flow to grid G_2. This means that the transconductance between grids 2 and 3 is negative, and a plot of the current and voltage

[1] Cledo Brunetti, The Transitron Oscillator, *Proc. IRE*, 27, p. 88, February, 1939.

on grid 2 under the conditions assumed will give a curve somewhat similar to that of Fig. 11-28 except that the current will never drop below the zero line. Thus this oscillator will perform in a manner similar to the dynatron with all the advantages of simplicity and of good frequency stability but without the uncertainty introduced by reliance on secondary emission.

A more practical circuit for the transitron oscillator is shown in Fig. 11-31. The condenser C_3 serves to pass on to grid 3 the alternating voltage of the tank circuit. If the time constant of the circuit consisting of C_3 and R_3 is long compared to the time of one cycle, the potential *across* condenser C_3 will remain nearly constant, thus approximating the constant battery voltage E_s of Fig. 11-30.

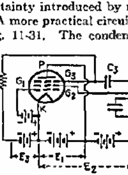


FIG. 11-31 Practical circuit of a negative transconductance oscillator

Beat-frequency Oscillator. It is sometimes desirable to have an audio oscillator which is continuously variable over a wide frequency spectrum. This result may be achieved by combining the output of two r-f oscillators which differ in frequency by the desired amount. As will be shown in Chap. 14 the output of the oscillators must be applied to a detector which will produce a signal having a frequency equal to the difference between the two

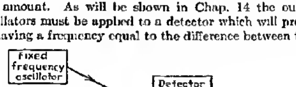


FIG. 11-32. Block diagram showing the principle of the beat-frequency oscillator.

oscillator frequencies. As indicated in the block diagram of Fig 11-32, one of the two oscillators is normally fixed in frequency while the other is variable over a small range. Thus if the fixed oscillator operates at, say, 100,000 cycles/sec, the variable oscillator might be operated over a range of 100,000 to 125,000 cycles/sec, producing a detector output with a range of 0 to 25,000 cycles/sec

on a single dial (the dial being mounted on the shaft of the condenser in the variable oscillator). Since the percentage frequency change of the variable oscillator is relatively small, its output can readily be maintained constant over the entire frequency range and the a-f output will then be constant in amplitude.

Multivibrators. A multivibrator is a relaxation type of oscillator, consisting essentially of a resistance-coupled amplifier in which part of the output is fed back to the input circuit. Oscillations are produced by the discharge of the coupling condensers through their associated grid leaks, the frequency being determined by the time constants of these circuits. Evidently such oscillations will be far from sinusoidal, a feature that makes the multivibrator an extremely valuable device for certain types of service, owing to the large number of harmonics present.

The circuit diagram of a typical multivibrator is given in Fig. 11-33, triode tubes being shown for the sake of simplicity although pentodes are more commonly used. Careful examination will show the similarity of this circuit to that of a resistance-coupled amplifier. The output of tube 1 is coupled to the input of tube 2 by means of the plate-load resistance R_{p_1} and the coupling condenser

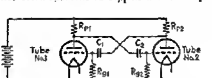


FIG. 11-33. Circuit of a multivibrator.

C_2. The output of tube 2 appears across R_{p_2}, but this is coupled back to the input of the first tube through coupling condenser C_1, thereby providing the feed-back feature characteristic of the multivibrator. The resistances R_{g_1} and R_{g_2} provide a conducting path from each grid to cathode, exactly as in the resistance-coupled amplifier (Chap. 9, page 265).

Suppose that a slight increase in plate current occurs in tube 1. The plate voltage of tube 1 is reduced; therefore, a more negative voltage appears on the grid of tube 2. This in turn reduces the plate current in tube 2, with a corresponding increase in its plate voltage. A more positive voltage is, therefore, induced on the grid of tube 1, causing a still further increase in its plate current.

With proper design this process is cumulative, and the plate current in tube 2 dies down almost instantly to zero while that in tube 1 builds up to a final value. This action is shown at time t_1, Fig. 11-34. From t_1 to t_2 the plate current of tube 2 remains at zero, but its grid gradually loses its charge through the resistance R_{g_2}. As soon as the grid potential becomes more positive than

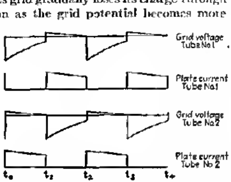


Fig. 11-34. Voltages and currents in the circuit of Fig. 11-33.

the cutoff voltage, plate current again flows, and the cycle repeats itself but with the action in the two tubes interchanged, the plate current in tube 1 now falling to zero while that in tube 2 rises to a maximum. This action repeats itself at intervals dependent upon the time of discharge of the two condensers C_1 and C_2. The discharge time of these two condensers need not be the same, the interval t_1t_2 differing from that of t_2t_3 if $C_1R_{g_1}$ is not equal to $C_2R_{g_2}$.[1]

The principal value of multivibrators lies in their irregular wave shape which is rich in harmonics. Wherever a large number of harmonics are desired, as in certain kinds of calibration work, these oscillators find a wide field of usefulness.

Controlled Multivibrators. The frequency of the multivibrator may be controlled from another source of alternating currents, such as a vacuum-tube oscillator. This is an invaluable aid in frequency measurements where a crystal-controlled oscillator may be used to control a multivibrator and so secure many accurately known frequencies throughout the spectrum, each harmonic of the multivibrator being known to the same degree of accuracy as the

[1] An excellent discussion of such dissymmetrical vibrators is given by L. M. Hull and J. K. Clapp, A Convenient Method of Referring Secondary Frequency Standards to a Standard Time Interval, *Proc. IRE*, 17, p. 252, February, 1929.

frequency of the crystal oscillator. For example, a 100-kc multivibrator under control of a crystal oscillator will produce harmonics at exact 100-kc intervals throughout the frequency spectrum. Similarly, a 10-kc multivibrator may be controlled by the same oscillator and so accurately produce frequencies throughout the frequency spectrum spaced 10 kc apart. Harmonics as high as the 150th have been utilized for calibration purposes.[1]

The circuit diagram of a controlled multivibrator is given in Fig. 11-35. A crystal oscillator is shown driving an isolating or buffer tube 3 the function of which is to prevent the abrupt changes of plate current in the multivibrator circuit from affecting the

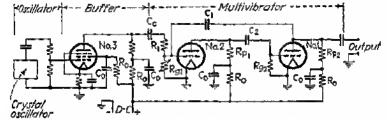


FIG. 11·35. Circuit of a multivibrator controlled by the output of a crystal oscillator.

frequency of the oscillator. A r-f pentode tube is used for this purpose to ensure complete isolation of the controlling oscillator. R_{p_1}, R_{p_2}, R_{g_1}, R_{g_2}, C_1, and C_2 in this diagram serve the same functions as in the circuit of Fig. 11-33. The resistances R_0 are voltage-dropping and filter resistances which prevent undesired coupling between the different tubes. The capacitances C_0 also prevent such coupling.

The output of the buffer amplifier inserts an additional (or synchronizing) voltage into the grid circuit of tube 1 of the multivibrator, shown by the curve of Fig. 11-36*a*. If the normal frequency of the vibrator is equal to that of the oscillator, the synchronizing voltage from tube 3 should have no effect, and time t_1 of Fig. 11-34 (when the grid of tube 1 goes positive) corresponds to the point of zero synchronizing voltage *a*, Fig. 11-36*a*. On the other hand if the normal frequency of the multivibrator is a

[1] A very thorough discussion of the use of multivibrators in frequency measurements including detailed circuits is given by F. E. Terman, "Measurements in Radio Engineering," McGraw-Hill Book Company, Inc., New York, 1935.

little lower than that of the controlling oscillator, the grid of tube 1 tends to become positive at a later point in the cycle of the synchronizing voltage, say at point *b*. However, the synchronizing voltage becomes increasingly positive following point *a* and, there-

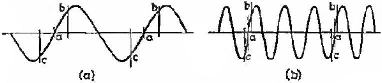


FIG. 11-36 Illustrating the possibility of synchronizing a multivibrator (*a*) at the frequency of the controlling oscillator, (*b*) at a submultiple of the oscillator frequency.

fore, causes tube 1 to conduct at a point between *a* and *b*, thus holding the multivibrator in step If the normal multivibrator frequency is a little higher than that of the controlling oscillator, the grid of tube 1 will tend to go positive at a point such as *c*, but the negative synchronizing voltage will prevent it from doing so until a time closer to *a* and so will again hold the multivibrator in step.

If the multivibrator frequency is very much less than that of the controlling signal, the synchronizing voltage will be unable to trip the grid of tube 1 until a later cycle as in Fig 11-36*b*, where the ratio of frequencies is 3:1. Here the grid of tube 1 becomes positive at the points *a*, or once in every three cycles of the synchronizing voltage, the normal multivibrator frequency being approximately one-third that of the controlling oscillator. Slight variations of the multivibrator frequency from the 3:1 ratio will cause the synchronizing voltage to operate to hold the vibrator in step in the same manner as in Fig. 11-36*a*, as indicated by the points *b* and *c* in the (*b*) figure. Proper adjustment of the multivibrator frequency and of the magnitude of the synchronizing voltage will permit synchronization at very high frequency ratios Hull and Clapp have successfully operated controlled multivibrators when the ratio of controlling to vibrator frequency was as high as 50:1, *i e*., a 50-kc oscillator can accurately control a 1-kc multivibrator.[1]

Parasitic Oscillations.[2] Oscillations that take place at other

[1] Hull and Clapp, *loc. cit.*

[2] For a more detailed discussion, see G. W Fyler, Parasites and Insta-

than the desired frequency, or outside the tank circuit, are termed *parasitic oscillations.* They may occur not only in oscillators but also in class A, B and C power amplifiers. Their presence is highly undesirable, since they may cause loss of efficiency, instability, flashover, shortened tube life, radiation at other than the desired frequency, and other undesirable effects. Their elimination is not an easy problem; and to predict their occurrence, before construction of the transmitter, is especially difficult. Thus the location and elimination of parasites are often found necessary after the radio transmitter has been installed, although much can be done in the original design to reduce the probability of their occurrence.

Parasites may be either of lower or of higher frequency than the normal frequency of the amplifier or oscillator in which they appear. Any circuit containing adequate energy storage and sufficient

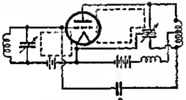


FIG. 11-37. Parasitic oscillations are often set up in the equivalent tuned-grid-tuned-plate circuit indicated by the dotted lines.

feedback of the proper phase will oscillate and may do so simultaneously with the main oscillations. Thus the effect of parasites is normally superimposed on the main output of the amplifier (or oscillator), making detection of such oscillations difficult. Undoubtedly many cases of poor performance have been due to unrecognized parasitic oscillations.

Parasitic oscillations may be set up in circuits that are similar to any of the common oscillator circuits, but probably most parasites are of the tuned-grid-tuned-plate type. An example is shown in Fig. 11-37 which shows the circuit of a neutralized class C amplifier. Parasitic oscillations may exist in the circuits indicated by dotted lines, constituting a tuned-grid-tuned-plate oscillator. At the frequency of this parasite the grid and plate

bility in Radio Transmitters, *Proc. IRE*, 23, p. 985, September, 1935; and J. S. Jackson, Analysis of Parasitic Oscillations in Radio Transmitters, *Radio News*, 35, p. 68, February, 1946.

tank capacitances present virtual short circuits, and the interelectrode capacitances together with the inductances of the grid and plate leads constitute the tank circuits for the parasite. The neutralizing condenser is evidently worthless at the parasitic frequency, since no voltage is built up across the plate tank circuit.

Elimination of this type of parasite may be accomplished in either of two ways: by inserting sufficient resistance in series with the parasitic current or by detuning the parasitic grid and plate circuits. In applying the former method a small resistance of perhaps 1 to 25 ohms is inserted in series with the grid or plate, preferably the former. This has but little effect on the normal operation of the tube, since it is not in series with the main oscillating circuit, but affects the parasite directly. Detuning is accomplished either by increasing the resonant frequency of the grid circuit or by decreasing that of the plate, since this will cause the plate circuit to present a capacitive reactance and thus introduce positive resistance into the grid circuit [see Eq. (9-7)]. The detuning process is carried out by shortening the grid leads to decrease their inductance or by actually inserting a small coil in series with the plate lead at a point close to the plate terminal of the tube.

Parasites may often be eliminated by the insertion of a small choke in the grid lead of the tube close to the grid terminal. This introduces appreciable impedance at the high parasitic frequencies but has little effect on the normal operation.

In circuits where chokes are used in series with the d-c supply to both plate and grid (e.g., L_g and L_s in Fig. 11-5), parasites may occur in which the frequency is determined by resonance between the chokes and the tank capacitances. These oscillations are also of the tuned-grid-tuned-plate type, and the remedy is evidently to make the natural frequency of the grid circuit higher than that of the plate. If the amplifier or oscillator is designed without a grid choke, the possibility of such oscillations is of course eliminated.

Another common source of parasitic oscillations is that of dynatron action in the grid circuit. The grid current may actually decrease with increasing grid voltage, throughout a certain voltage range, owing to secondary emission. The a-c grid resistance is then negative over that range, and parasitic oscillations may be set up by dynatron action.

Relaxation oscillations are sometimes encountered, similar to the action of multivibrators (page 481). These have already been referred to as "intermittent oscillations" (page 461) and may occur at either low or high frequencies.

A thorough study of parasitic oscillations is very complex, since parasites occur in a great variety of circuits. In general one should at least suspect the possibility of parasitic oscillations whenever the performance of a class B or C amplifier or an oscillator is abnormal.[1]

Ultra-high-frequency Oscillators. At very high frequencies, as above about 500 Mc, conventional oscillator circuits and tubes are unsatisfactory owing to the effect of electrode and circuit capacitances and the transit time of the electron in traversing the tube. For these frequencies a whole new group of tubes, including the magnetron and klystron, have been designed, using entirely different circuits. A detailed discussion of these tubes and circuits and of their performance is beyond the scope of this book.

Problems

11-1. Determine the constants of the tank circuit of a Hartley oscillator for the following: $E_b = 2000$ volts, $E_p = 1280$ volts, $E_g = 450$ volts, power delivered to the tank circuit = 375 watts, $f = 1250$ kc. Carry out the design for (a) $Q = 12.5$ and (b) $Q = 75$.

11-2. Repeat Prob. 11-1 for a Colpitts oscillator.

11-3. In the circuit of Fig. 11-16, C_1 and C_2 are twin variable condensers mounted on a common shaft, and adjustable to give from 0.001 μf max to 0.0001 μf min including wiring and other stray circuit capacitances. R_1 and R_2 are to be kept equal but are adjustable in steps by means of a rotating switch. (a) How many steps are required if the oscillator is to cover the frequency range from 20 to 200,000 cycles/sec? (b) What must be the values of R_1 and R_2 for each step?

11-4. A certain crystal has the following constants: $L = 0.29$ henry, $C = 0.069$ $\mu\mu f$, $R = 2900$ ohms. (a) What is the Q of the crystal? (b) If the shunting capacitance of the holder together with the input capacitance of the tube is 7 $\mu\mu f$, at what frequency will the crystal and associated circuit be resonant? (c) What is the percentage of change in the resonant frequency if the shunting capacitance is reduced to 5 $\mu\mu f$?

11-5. The tube to which the curves of Fig. 3-40 apply is used in the dynatron circuit of Fig. 11-27, where $E_b = 30$ volts, $E_{c2} = 100$ volts, $E_{c1} = 0$, $L = 20\,\mu h$, $C = 0.0002\,\mu f$. (a) What is the maximum resistance which may be used in series with the coil L in the tank circuit (Q of the coil is 150) and still permit the circuit to oscillate? (b) What is the approximate tube output with this resistance?

[1] For a discussion of general methods of locating parasites and for their removal, see Fyler, *loc. cit.*

CHAPTER 12

POWER-SERIES ANALYSIS OF VACUUM-TUBE PERFORMANCE[1]

Analysis of vacuum-tube performance has been carried out in the preceding chapters of this book either by assuming that the current-voltage characteristics were sufficiently linear to permit the use of rms values and vectors at a single frequency or by using step-by-step analysis of the instantaneous values of the current flowing. For many purposes a more general analytical approach is desirable, one that is applicable regardless of the wave shape of the currents and potentials. A common mathematical tool for handling this type of problem is the power series, and this proves to be a useful method of analyzing the performance of vacuum tubes in the general case

Power-series Expansion of the Plate Current. The equation of the plate current in a triode (or in a multigrid tube if all the electrodes except the control grid and plate are maintained at a constant potential) was given in Eq. (3-8), page 60, and is reproduced herewith:

$$i_b = f(e_b, e_c) \tag{12-1}$$

The plate and grid voltages and the plate current of the tube each consist of a zero-signal direct component and a component due to the a-c excitation on the grid. Thus we may write $e_b = E_{b0} + e_{p0}$, $e_c = E_c + e_g$, and $i_b = I_{b0} - i_{p0}$.[2] Furthermore, the incremental plate voltage e_{p0} is equal to the product of the incre-

[1] This chapter may be omitted by those who are not interested in the more advanced treatments of the two following chapters

[2] The subscript $p0$ is used with the plate voltage and current rather than the usual p of Eqs. (9-75), (9-74), and (9 73), since a rectified direct component of current may appear in the plate circuit of the tube under certain conditions, increasing the average, or direct, plate current from I_{b0} to I_b and the voltage from E_{b0} to E_b. Thus i_{p0}, which may be referred to as the *incremental* plate current, is the difference between I_{b0} and i_b in which i_{p0}

mental plate current and the load impedance. For the present this impedance is assumed to be purely resistive, and we may write[1]

$$e_{p0} = i_{p0}R_L \tag{12-2}$$

From the foregoing relations we may rewrite Eq. (12-1) as

$$i_b = I_{b0} - i_{p0} = f(E_{bb} + i_{p0}R_L, E_c + e_g) \tag{12-3}$$

Equation (12-3) may be expanded by means of Taylor's theorem to give[2]

$$i_b = I_{b0} - i_{p0} = I_{b0} + i_{p0}R_L\frac{\partial i_b}{\partial e_b} + e_g\frac{\partial i_b}{\partial e_c} + \frac{i_{p0}^2R_L^2}{2}\frac{\partial^2 i_b}{\partial e_b^2} + i_{p0}R_Le_g\frac{\partial^2 i_b}{\partial e_b\,\partial e_c} + \frac{e_g^2}{2}\frac{\partial^2 i_b}{\partial e_c^2} + \cdots \tag{12-4}$$

in which the derivatives are to be evaluated at the point E_{b0}, E_c.

---

may contain a direct component, whereas i_p, the alternating plate current, is the difference between I_b and i_b and contains no direct component. (See Appendix A where these symbols are illustrated in a figure on page 608 as applied to the plate current. The application to the plate voltage is similar.)

[1] When the load is other than resistive, its impedance will, in general, vary with frequency and will therefore be different for each harmonic of the plate current. The analysis must then be generalized in the manner presented in a later section of this chapter, p. 500.

[2] In textbooks on calculus Taylor's theorem is given in the following form when applied to the expansion of a function of two variables:

$$f(x, y) = f(x_0 + h, y_0 + k) = f(x_0, y_0) + \left(h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y}\right)f(x, y)\Big|_{\substack{x=x_0\\y=y_0}} + \frac{1}{2!}\left(h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y}\right)^2 f(x, y)\Big|_{\substack{x=x_0\\y=y_0}} + \frac{1}{3!}\left(h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y}\right)^3 f(x, y)\Big|_{\substack{x=x_0\\y=y_0}} + \cdots$$

where $\left(h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y}\right)^n$ is to be expanded by the binomial theorem as if the expression within the parentheses were two terms of a binomial, and the resulting terms applied separately to $f(x, y)$. The resulting derivatives $\partial f(x, y)/\partial x$, etc., are to be evaluated at the point x_0, y_0.

In the expansion of Eq. (12-3) $x = e_b$, $y = e_c$, $x_0 = E_{b0}$, $y_0 = E_c$, $h = e_{p0} = i_{p0}R_L$, $k = e_g$, $f(x, y) = f(e_b, e_c) = i_b$ and $f(x_0, y_0) = f(E_{b0}, E_c) = I_{b0}$. Evidently then $\frac{\partial f(x, y)}{\partial x} = \frac{\partial i_b}{\partial e_b}$ and must be evaluated at the point E_{b0}, E_c. Other derivatives are treated similarly.

This is a power series and, when solved explicitly for i_{pa}, will be of the form[1]

$$i_{pa} = a_1 e_g + a_2 e_g^2 + a_3 e_g^3 + \cdots \tag{12-5}^*$$

where a_1, a_2 $\cdots$ are constants.

The constants a_1, a_2 $\cdots$ may be evaluated by substituting Eq. (12-5) into Eq. (12-4) and equating the multiplying factors of e_g, e_g^2 $\cdots$ on each side of the equation.[2] The series converges quite rapidly, and for many purposes all terms of higher order than the second may be neglected. Under such an assumption the resulting equation will yield for the first two constants

$$a_1 = -\frac{\dfrac{\partial i_b}{\partial e_c}}{1 + R_L \dfrac{\partial i_b}{\partial e_b}} \tag{12-6}^*$$

$$a_2 = -\frac{R_L^2 \left(\dfrac{\partial i_b}{\partial e_c}\right)^2 \dfrac{\partial^2 i_b}{\partial e_b^2}}{2\left(1 + R_L \dfrac{\partial i_b}{\partial e_b}\right)^3} + \frac{R_L \dfrac{\partial i_b}{\partial e_c} \dfrac{\partial^2 i_b}{\partial e_b \partial e_c}}{\left(1 + R_L \dfrac{\partial i_b}{\partial e_b}\right)^2} - \frac{\dfrac{\partial^2 i_b}{\partial e_c^2}}{2\left(1 + R_L \dfrac{\partial i_b}{\partial e_b}\right)} \tag{12-7}^*$$

Equations (12-6) and (12-7) may be further simplified by replacing the first-order derivatives with their equivalent values from Eqs. (3-24b) and (3-24c); thus $\partial i_b/\partial e_b = 1/r_p$ and $\partial i_b/\partial e_c = g_m$. With these substitutions and the aid of Eq. (3-13), a_1 becomes[3]

[1] That Eq. (12-5) is actually a solution of Eq. (12-4) may be proved by substituting the former into the latter and attempting to solve for the constants a_1, a_2, $\cdots$. If Eq. (12-5) is a solution, its substitution should yield relations for a_1, a_2, $\cdots$ that involve only constant terms. This procedure is carried out in the succeeding demonstration. (For the meaning of * after an equation number, see Preface to the First Edition.)

[2] For further detail see F. B. Llewellyn, Operation of Thermionic Vacuum Tube Circuits, *Bell System Tech. J.*, 5, p. 433, July, 1926. The signs in the equations of the constants developed here are opposite to those of Llewellyn's owing to the minus sign between I_{bo} and i_{pa}.

[3] The expression given here for a_1 and that of Eq. (12-16) for a_2, it must be remembered, are based on the assumption of resistance load; i.e., the load presents the same impedance to all components of the plate current, regardless of their frequency. When it contains some reactive element, a_1 and a_2 are given by Eqs. (12-46) and (12-47), p. 501, and Eq. (12-5) must be used in the manner described in the section beginning on p. 500.

$$a_1 = -\frac{g_m r_p}{r_p + R_L} \tag{12-8}^*$$

or

$$a_1 = -\frac{\mu}{r_p + R_L} \tag{12-9}^*$$

where μ, g_m, and r_p are to be evaluated at the point (E_{b0}, E_c). For pentodes R_L is normally very much smaller than r_p, and we may write

$$a_{1(\text{pentodes})} = -\frac{\mu}{r_p} = -g_m \qquad \text{(approx)} \tag{12-10}^*$$

We may also write

$$\frac{\partial^2 i_b}{\partial e_b^2} = \frac{\partial}{\partial e_b}\left(\frac{\partial i_b}{\partial e_b}\right) = \frac{\partial \dfrac{1}{r_p}}{\partial e_b} = -\frac{1}{r_p^2}\frac{\partial r_p}{\partial e_b} \tag{12-11}$$

and

$$\frac{\partial^2 i_b}{\partial e_b\,\partial e_c} = \frac{\partial}{\partial e_b}\left(\frac{\partial i_b}{\partial e_c}\right) = \frac{\partial}{\partial e_b}\left(\frac{\mu}{r_p}\right) = \frac{1}{r_p}\frac{\partial \mu}{\partial e_b} - \frac{\mu}{r_p^2}\frac{\partial r_p}{\partial e_b} \tag{12-12}$$

From Eqs. (3-13) and (3-24) we may write

$$\mu = g_m r_p = \frac{\partial i_b/\partial e_c}{\partial i_b/\partial e_b}$$

or

$$\frac{\partial i_b}{\partial e_c} = \mu\,\frac{\partial i_b}{\partial e_b} \tag{12-13}$$

Then,

$$\frac{\partial^2 i_b}{\partial e_c^2} = \frac{\partial}{\partial e_c}\left(\frac{\partial i_b}{\partial e_c}\right) = \frac{\partial}{\partial e_c}\left(\mu\,\frac{\partial i_b}{\partial e_b}\right) = \mu\,\frac{\partial^2 i_b}{\partial e_b\,\partial e_c} + \frac{\partial i_b}{\partial e_b}\frac{\partial \mu}{\partial e_c} \tag{12-14}$$

Equation (12-14) may be simplified by means of Eq. (12-12), at the same time replacing $\partial i_b/\partial e_b$ with $1/r_p$ and using g_m in place of μ/r_p, giving

$$\frac{\partial^2 i_b}{\partial e_c^2} = g_m\,\frac{\partial \mu}{\partial e_b} - g_m^{\,2}\,\frac{\partial r_p}{\partial e_b} + \frac{1}{r_p}\frac{\partial \mu}{\partial e_c} \tag{12-15}$$

Substituting the foregoing relations into Eq. (12-7) and letting $\partial r_p/\partial e_b = r_p'$, $\partial \mu/\partial e_b = \mu_b'$ and $\partial \mu/\partial e_c = \mu_c'$ gives

$$a_2 = \frac{\mu^2 r_p r_p' - \mu\mu_b'(r_p^2 - R_L^2) - \mu_c'(r_p + R_L)^2}{2(r_p + R_L)^3} \qquad (12\text{-}16)^*$$

where μ, r_p, r_p', μ_b', and μ_c' are to be evaluated at the point (E_{bb}, E_c).

In a triode μ is reasonably constant, and a fair degree of approximation is obtained by letting $\mu_b' = 0$ and $\mu_c' = 0$ to give

$$a_{2\,(\text{triode})} = \frac{\mu^2 r_p r_p'}{2(r_p + R_L)^3} \qquad \text{(approx)} \qquad (12\text{-}17)^*$$

This approximation is entirely unsatisfactory for pentodes and Llewellyn has shown that, even in triodes, it may produce appreciable error.

For pentode and beam tubes we may write $R_L \ll r_p$ and Eq. (12-16) may be simplified by neglecting R_L in each of the three parentheses to give, when simplified,

$$a_{2\,(\text{pentode})} = \frac{1}{2}\left(\frac{\mu^2 r_p'}{r_p^2} - \frac{\mu\mu_b'}{r_p} - \frac{\mu_c'}{r_p}\right) \qquad \text{(approx)} \qquad (12\text{-}18)^*$$

For pentodes it is better to use g_m rather than μ. With the aid of Eq. (3-13), page 62, we may rewrite Eq. (12-18) as

$$a_{2\,(\text{pentode})} = \frac{1}{2}\left(g_m^2 r_p' - g_m\mu_b' - \frac{\mu_c'}{r_p}\right) \qquad \text{(approx)} \qquad (12\text{-}19)^*$$

A better approach to this problem for pentode and beam tubes is to obtain a_2 in terms of derivatives of g_m instead of μ. To do so, let us write

$$\frac{\partial^2 i_b}{\partial e_c^2} = \frac{\partial}{\partial e_c}\left(\frac{\partial i_b}{\partial e_c}\right) = \frac{\partial g_m}{\partial e_c} = g_{mc}' \qquad (12\text{-}20)$$

$$\frac{\partial^2 i_b}{\partial e_b\,\partial e_c} = \frac{\partial}{\partial e_b}\left(\frac{\partial i_b}{\partial e_c}\right) = \frac{\partial g_m}{\partial e_b} = g_{mb}' \qquad (12\text{-}21)$$

Substituting Eqs. (12-20) and (12-21) into Eq. (12-7) gives

$$a_2 = \frac{R_L^2 g_m^2 r_p r_p' + 2R_L g_m r_p^2 (r_p + R_L) g_{mb}' - (r_p + R_L)^2 r_p g_{mc}'}{2(r_p + R_L)^3} \qquad (12\text{-}22)^*$$

For pentodes R_L is normally very much less than r_p so that $(r_p + R_L)$ is very nearly equal to r_p. Equation (12-22) may, therefore, be written

$$a_2 = \frac{R_L^2 g_m^2 r_p'}{2r_p^2} + R_L g_m g'_{mb} - \frac{g'_{mc}}{2} \qquad \text{(approx)} \qquad (12\text{-}23)$$

The first two terms in Eq. (12-23) are very small compared to the third[1] so that we may write

$$a_{2\,(\text{pentodes})} = -\frac{g'_{mc}}{2} = -\frac{1}{2}\frac{\partial g_m}{\partial e_c} \qquad \text{(approx)} \qquad (12\text{-}24)^*$$

[1] This may be demonstrated by showing that the ratio of the third term to the second is large compared to unity and that the ratio of the second term to the first is greater than one, or, in mathematical form, that

$$\frac{\text{Third term}}{\text{Second term}} = \frac{g'_{mc}}{2R_L g_m g'_{mb}} \gg 1 \qquad (1)$$

and

$$\frac{\text{Second term}}{\text{First term}} = \frac{2r_p^2 g'_{mb}}{R_L g_m r'_p} > 1 \qquad (2)$$

To demonstrate (1) we may first write from Eq. (12-20)

$$g'_{mc} = \frac{\partial}{\partial e_c}\left(\frac{\partial i_b}{\partial e_c}\right) \qquad (3)$$

Substituting Eq. (12-13) into (3) we may write

$$g'_{mc} = \frac{\partial}{\partial e_c}\left(\mu \frac{\partial i_b}{\partial e_b}\right) = \mu \frac{\partial^2 i_b}{\partial e_b \partial e_c} + \frac{\partial i_b}{\partial e_b}\frac{\partial \mu}{\partial e_c} \qquad (4)$$

Using Eqs. (12-21) and (3-24b) this becomes

$$g'_{mc} = \mu \frac{\partial g_m}{\partial e_b} + \frac{1}{r_p}\frac{\partial \mu}{\partial e_c} \qquad (5)$$

or

$$g'_{mc} = r_p g_m g'_{mb} + \frac{\mu'_c}{r_p} \qquad (6)$$

Substituting this relation into (1) gives

$$\frac{\text{Third term}}{\text{Second term}} = \frac{r_p}{2R_L} + \frac{\mu'_c}{2R_L r_p g_m g'_{mb}} \qquad (7)$$

Since $R_L \ll r_p$, the first term alone is large compared to unity, so that the inequality expressed in (1) is proved without determining the size of the second term of Eq. (7).

To demonstrate (2) we may write from Eq. (12-21)

While the first and second order terms are the most important in Eq. (12-5), the third-order terms are sometimes of interest. These can be found by extending the preceding methods to include the terms containing e_g^3. For pentode and beam tubes the resulting term may be simplified[1] to merely

$$a_{3(\text{pentodes})} = \frac{1}{3!}\frac{\partial^2 g_m}{\partial e_c^2} \qquad \text{(approx)} \qquad (12\text{-}25)^*$$

We may now summarize the foregoing by writing the two following equations for the plate current in a triode and in a pentode or beam tube. These are obtained by substituting the appropriate values of i_{p0} into the relation $i_b = I_{b0} - i_{p0}$

For triodes (if μ may be considered constant),

$$i_b = I_{b0} + \frac{\mu e_g}{r_p + R_L} - \frac{\mu^2 r_p r_p' e_g^2}{2(r_p + R_L)^3} + \cdots \qquad (12\text{-}26)^*$$

For pentodes (assuming $R_L \ll r_p$),

$$i_b = I_{b0} + g_m e_g + \frac{1}{2!}\frac{\partial g_m}{\partial e_c} e_g^2 + \frac{1}{3!}\frac{\partial^2 g_m}{\partial e_c^2} e_g^3 + \cdots \qquad (12\text{-}27)^*$$

$$g_{mb}' = \frac{\partial}{\partial e_b}\left(\frac{\partial i_b}{\partial e_c}\right) \qquad (8)$$

Substituting Eq. (12-13) into (8) gives

$$g_{mb}' = \frac{\partial}{\partial e_b}\left(\mu \frac{\partial i_b}{\partial e_b}\right) = \mu \frac{\partial^2 i_b}{\partial e_b^2} + \frac{\partial i_b}{\partial e_b}\frac{\partial \mu}{\partial e_b} \qquad (9)$$

Substituting Eqs. (12-11) and (3-24b) into (9) gives

$$g_{mb}' = -\frac{\mu}{r_p^2} r_p' + \frac{1}{r_p}\mu_b' \qquad (10)$$

Substituting (10) into (2) gives

$$\frac{\text{Second term}}{\text{First term}} = -\left(\frac{2\mu}{R_L g_m} - \frac{2 r_p \mu_b'}{R_L g_m r_p'}\right) \qquad (11)$$

The first term in Eq. (11) may be written $2\mu/R_L g_m = 2r_p g_m/R_L g_m = 2r_p/R_L$ which, since $R_L \ll r_p$, is large compared to unity. The second term in Eq (11) is actually additive to the first, not subtractive, since r_p' is numerically negative (plate resistance *decreases* as e_b *increases*) so that the inequality expressed in (2) is proved. Thus the simplification of Eq. (12-24) is legitimate.

[1] For further information on terms of higher order than the second, see Llewellyn, *loc. cit.*

Amplitude Distortion. If a sinusoidal emf is impressed on the grid of an amplifier tube, we may write

$$e_g = E_{gm} \sin \omega t \tag{12-28}$$

which, substituted in Eq. (12-5), gives

$$i_{p0} = a_1 E_{gm} \sin \omega t + a_2 E_{gm}^2 \sin^2 \omega t + a_3 E_{gm}^3 \sin^3 \omega t + \cdots \tag{12-29}$$

but

$$\sin^2 \omega t = \frac{1 - \cos 2\,\omega t}{2} \tag{12-30}$$

$$\sin^3 \omega t = \frac{3 \sin \omega t - \sin 3\,\omega t}{4} \tag{12-31}$$

Substituting Eqs. (12-30) and (12-31) into (12-29) and rearranging terms gives

$$i_{p0} = \frac{a_2 E_{gm}^2}{2} + \left(a_1 E_{gm} + \frac{3a_3 E_{gm}^3}{4}\right) \sin \omega t - \frac{a_2 E_{gm}^2 \cos 2\,\omega t}{2} - \frac{a_3 E_{gm}^3 \sin 3\,\omega t}{4} + \cdots \tag{12-32}$$

Equation (12-32) shows that the output of a vacuum tube operating on the nonlinear portion of its characteristic curve and having a sine wave of emf applied will, in general, contain an infinite number of harmonics of the impressed frequency. Furthermore, since $a_1 > a_2 > a_3 \cdots$, the higher harmonics will be of decreasing magnitude. It is seldom necessary to consider anything higher than the fifth harmonic; in many cases the second is sufficiently high.

Equation (12-32) also shows that the current i_{p0} contains a direct component $a_2 E_{gm}^2/2$; this was the reason why the subscript $p0$ was used instead of p. As should also be noted, the magnitude of this direct component is equal to that of the second harmonic;[1] consequently the increase in direct plate current, which follows the application of a-c grid excitation, is a direct indication of the magnitude of the second-harmonic distortion present.

[1] This is exactly true only if no components of higher order than the third are present. Higher order even harmonics will be accompanied by a further change in the direct current, although such change is normally

It may also be seen from Eq. (12-32) that distortion will be absent only when $a_2, a_3 \cdots = 0$, since only then will the harmonic terms disappear (*i.e.*, terms of two, three, etc., times the fundamental frequency of $\omega/2\pi$). This will be the case whenever the second-order derivatives of plate current are zero, *i.e.*, when the static characteristic curves are straight.[1] Thus we again come to the conclusion that distortionless amplification may be secured only by limiting the operation of the amplifier to the substantially straight portions of the tube characteristic curves. Furthermore, if $a_2, a_3 \cdots$ are zero, Eq. (12-5) reduces to $i_{p0} = a_1 e_p$, and substitution of Eq. (12-9) gives

$$i_{p0} = \frac{-\mu e_g}{r_p + R_L} \tag{12-33}$$

Since no direct component is present in Eq. (12-33), $i_{p0} = i_p$ and the foregoing equation may be written

$$-\mu e_g = i_p(r_p + R_L) \tag{12-34}$$

which is the general equation for the vacuum tube with resistance load, under distortionless conditions, previously developed on page 63 and given as Eq. (3-19). Thus this important equation has again been developed, this time by more rigorous methods.

If the voltage applied to the grid of an amplifier tube consists of two or more simultaneously impressed sine waves (as in the amplification of speech or music), many more distortion terms will be present than are indicated by Eq. 12-32). If Eq. (12-28) is written as

$$e_g = E_{g1m} \sin \omega t + E_{g2m} \sin qt + E_{g3m} \sin pt + \cdots \tag{12-35}$$

(where $\omega/2\pi$, $q/2\pi$, and $p/2\pi$ are different audio frequencies) and this value of e_g is inserted in Eq. (12-5), the plate-current com-

quite small. Also, if the load impedance presented to the d-c component is different from that for the a-c components, as in a transformer-coupled amplifier, the ratio of the magnitude of the d-c component to that of the second harmonic is larger than indicated in Eq. (12-32). (See discussion beginning on p. 500.)

[1] Inspection of Eq. (12-7) shows that for a_2 to be zero all second-order derivatives must be zero. If higher order constants are evaluated in the same manner as was a_2, it will be found that a_n will be zero if $\partial^n i_b/\partial e_b{}^n$, $\partial^n i_b/\partial e_c{}^n \cdots$ are zero. Since all higher order derivatives must be zero if the second-order terms are zero for all values of e_b and e_c, the constants a_2, $a_3 \cdots$ will all be zero if $a_2 = 0$ for all values of e_b and e_c.

ponents will include not only double, triple, etc., the impressed frequencies but also frequencies equal to the sum and difference of each pair, the sum and difference of one and twice another, etc.[1] Thus amplitude distortion is characterized by the introduction of a large number of new frequency components in the output of the amplifier.

Power-series Expansion of the Plate Current. Diodes. It is possible to develop equations for the diode which are exactly like those for the triode and pentode if we assume the diode to be the equivalent of a triode with $\mu = 1$. Treatment of the diode in this manner seems desirable in the interests of simplicity, as a single set of equations will then serve all types of tubes.

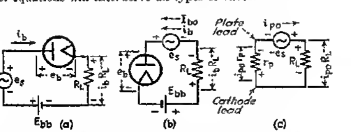


FIG. 12-1. Development of an equivalent circuit for a diode to be compared to that of Figs. 3-22 and 3-24 for a triode.

Figure 12-1a shows a diode tube with a load resistance R_L, a source of impressed alternating voltage e_s, and a source of direct voltage E_{bb}.[2] The equation of the plate current flowing is given by the equation of the static characteristic curve $i_b = f(e_b)$, or

$$i_b = f(e_b) = f(E_{bb} + e_s - i_b R_L) \qquad (12\text{-}36)$$

the assumed positive directions of the alternating emfs being indicated on the figure by plus and minus signs.[3]

In Fig. 12-1b the circuit has been rearranged for comparison with the simple triode circuit of Fig. 3-22 (page 58) from which it differs only in that the excitation voltage is inserted into the plate

[1] This phenomenon is treated in more detail under Square-law Modulation, pp. 529–531.

[2] Diodes are usually operated without a d-c source, but the d-c supply is included here to make the development general. The solution will apply to the case of no d-c supply by merely making $E_{bb} = 0$.

[3] This use of + and − signs is explained in Appendix E.

circuit, instead of into the grid circuit as in Fig. 3-22. It is, therefore, possible to replace the circuit of Fig. 12-1b with an equivalent circuit for the incremental components only, similar to that of the triode shown in Fig. 3-24 (page 64). This has been done in Fig. 12-1c where r_p is the a-c resistance of the diode (see discussion on page 47). To complete the analogy with the triode, the incremental component of plate current is assumed to flow out of the plate terminal on its positive half cycle, as indicated by the arrow.[1] Since the flow of incremental plate current is assumed to be opposite in direction to that of the direct component, it is desirable to reverse the polarity of the alternating voltage, therefore the equivalent alternating voltage is shown as $-e_s$, where e_s has the polarity indicated in Fig. 12-1a. This is, of course, equivalent to the voltage $-\mu e_g$ of the triode circuit (Fig. 3-24).

This comparison of the diode with the triode presents one apparent ambiguity, *viz.*, the plate voltage of the triode, as shown in Fig. 3-24, is *the voltage across the load*, whereas this is *not* the case with the diode. This apparently makes the diode plate voltage negative (*i.e.*, $e_{p\delta} = -i_{p0}r_p$). If, however, the load voltage of Fig. 12-1c is considered as the equivalent of the plate voltage of Fig. 3-24, the circuit is identical with that of the triode, and the equations previously developed for the triode should be applicable to the diode by letting $\mu = 1$. As a matter of fact, this ambiguity in the plate-voltage notation should cause no trouble, as there is little occasion to be interested in the voltage across the diode except to consider it as the drop through the a-c resistance r_p, and the greater simplicity of development made possible by comparing the diode with the triode through the notation of Fig. 12-1 is highly desirable.

From the notation of Fig. 12-1 we may write

$$i_b = I_{b0} - i_{p0} \tag{12-37}$$

which corresponds to the equation of i_b for a triode as given on page 488. Substituting Eq. (12-37) into Eq. (12-36) gives

$$i_b = I_{b0} - i_{p0} = f(E_{bb} + e_s - I_{b0}R_L + i_{p0}R_L) \tag{12-38}$$

Since $E_{bb} - I_{b0}R_L = E_{b0}$ (*i.e.*, the direct plate voltage is equal to

[1] The symbol i_{p0} is used rather than i_p for the reason outlined in the footnote on p. 488.

the supply voltage minus the d-c drop through the load), Eq. (12-38) may be further simplified to

$$i_b = I_{b0} - i_{p0} = f(E_{b0} + e_s + i_{p0}R_L) \tag{12-39}$$

This equation may be expanded by Taylor's theorem to give[1]

$$i_b = I_{b0} - i_{p0} = I_{b0} + e_s \frac{di_b}{de_b} + i_{p0} R_L \frac{di_b}{de_b} + \frac{e_s^2}{2} \frac{d^2 i_b}{de_b^2} + e_s i_{p0} R_L \frac{d^2 i_b}{de_b^2} + \frac{i_{p0}^2 R_L^2}{2} \frac{d^2 i_b}{de_b^2} + \cdots \tag{12-40}$$

As in the case of the triode, if Eq. (12-40) is solved explicitly for i_{p0}, the resulting equation is of the form

$$i_{p0} = b_1 e_s + b_2 e_s^2 + b_3 e_s^3 + \cdots \tag{12-41}$$

Inserting Eq. (12-41) into Eq. (12-40) permits solution of the constants b_1, $b_2 \cdots$ as described on page 490. Neglecting all terms of higher order than the second gives for b_1 and b_2

$$b_1 = -\frac{\dfrac{di_b}{de_b}}{1 + R_L \dfrac{di_b}{de_b}} \tag{12-42}$$

$$b_2 = -\frac{R_L^2 \left(\dfrac{di_b}{de_b}\right)^2 \dfrac{d^2 i_b}{de_b^2}}{2\left(1 + R_L \dfrac{di_b}{de_b}\right)^3} + \frac{R_L \dfrac{di_b}{de_b} \dfrac{d^2 i_b}{de_b^2}}{\left(1 + R_L \dfrac{di_b}{de_b}\right)^2} - \frac{\dfrac{d^2 i_b}{de_b^2}}{2\left(1 + R_L \dfrac{di_b}{de_b}\right)} \tag{12-43}$$

the derivations being evaluated at the point (I_{b0}, E_{b0}).

[1] Taylor's theorem was given in a footnote on p. 489, as applied to the expansion of a function of two variables. In the expansion of Eq. (12-39), $e_s + i_{p0}R_L$ may be considered a single variable, and Taylor's theorem as applied to the expansion of a single variable is

$$f(x) = f(x_0 + h) = f(x_0) + h \frac{df(x)}{dx}\bigg|_{x=x_0} + \frac{h^2}{2!} \frac{d^2 f(x)}{dx^2}\bigg|_{x=x_0} + \frac{h^3}{3!} \frac{d^3 f(x)}{dx^3}\bigg|_{x=x_0} + \cdots$$

Equations (12-42) and (12-43) may be further simplified by noting from page 47 that

$$\frac{di_b}{de_b} = \frac{1}{r_p}$$

and

$$\frac{d^2 i_b}{de_b^2} = \frac{d}{de_b}\left(\frac{di_b}{de_b}\right) = \frac{d(1/r_p)}{de_b} = -\frac{1}{r_p^2}\frac{dr_p}{de_b}$$

This gives

$$b_1 = -\frac{1}{r_p + R_L} \tag{12-44}$$

and

$$b_2 = \frac{r_p r_p'}{2(r_p + R_L)^3} \tag{12-45}$$

where

$$r_p' = \frac{dr_p}{de_b}$$

Comparison of Eqs. (12-44) and (12-45) with Eqs (12-9) and (12-16) or (12-17) shows that the equations for a_1 and a_2 may be used for b_1 and b_2, respectively, by letting $\mu = 1$. For the balance of this book, therefore, the triode equations will be used for both diodes and triodes.

Effect of Load Impedances Which Are a Function of Frequency. Triodes. Equations (12-9) and (12-16) were developed assuming a load resistance R_L which is independent of frequency whereas actually the load impedance may vary with frequency, producing a different effect on each component of the plate current. Equations (12-9) and (12-16) must therefore be modified to fit the general case. If the expansion of Eq. (12-1) is repeated for the general case of a load impedance Z_L, it will be found that Eq. (12-5) is still satisfactory, but solution of the first two constants gives[1]

The derivatives $df(x)/dx$, etc., must be evaluated at the point x_0.

In the expansion of Eq (12-39) $x = e_b$, $x_0 = E_{bo}$, $h = e_s + i_{pa}R_L$, $f(x) = f(e_b) = i_b$, and $f(x_0) = f(E_{bo}) = I_{bo}$. Evidently then $df(x)/dx = di_b/de_b$ and must be evaluated at the point $e_b = E_{bo}$. Other derivatives are treated similarly.

[1] The derivation of the constants for a general impedance is too extensive a development for presentation here, and the reader is referred to Llewellyn, *loc. cit.* It should be noted that the constants given here are of opposite

$$a_{1(m)} \doteq -\frac{\mu}{|r_p + Z_{L(m)}|} \tag{12-46}^*$$

and

$$a_{2(m\pm n)} = \frac{\mu^2 r_p r_p' - \mu\mu_b' |r_p{}^2 - Z_{L(m)}Z_{L(n)}| - \mu_s' |(r_p + Z_{L(m)})(r_p + Z_{L(n)})|}{2|(r_p + Z_{L(m)})(r_p + Z_{L(n)})(r_p + Z_{L(m\pm n)})|} \tag{12-47}^*$$

where r_p', μ_b', and μ_s' have the same significance as in Eq. (12-16). As previously stated, the second and third terms in the numerator may often be neglected for triodes, are zero for diodes, but must be included for multigrid tubes (although a simpler solution for multigrid tubes is given in a later section).

The subscripts m and n are the angular frequencies[1] of the particular impressed signals which enter into the term of which a_1 or a_2 is a part. Thus Eq. (12-32), for example, was developed from Eq. (12-29), so that the angular frequency of the signal which produced the component $a_1 E g_m \sin \omega t$ in Eq. (12-32) is ω. For this equation, then, a_1 would be evaluated from Eq. (12-46) by writing $m = \omega$ and the impedance $Z_{L(m)}$ is the impedance of the load circuit at the angular frequency $m = \omega$. Similarly those terms of Eq. (12-32) which contain a_2 resulted from the expansion of the second term of Eq. (12-29) and are therefore derived from $\sin^2 \omega t$ or $(\sin \omega t)(\sin \omega t)$. The frequency of the sine function in the first of these parentheses determines m and the second determines n, so that in this case $m = n = \omega$. The first term in Eq. (12-32) was derived by using the first term in Eq. (12-30), *i.e.*, ½. The angular frequency of this term is $m - n = \omega - \omega = 0$; therefore, a_2 in the first term of Eq. (12-32) is really $a_{2(\omega-\omega)}$ and may be evaluated from Eq. (12-47) where $Z_{L(m\pm n)}$ becomes $Z_{L(\omega-\omega)}$ and is equal to the impedance of the load circuit at zero frequency, while $Z_{L(m)}$ and $Z_{L(n)}$ become $Z_{L(\omega)}$ and are equal to the impedance of the load at the angular frequency ω.

---

sign to those of Llewellyn, because of the use of the negative sign in $i_b = I_{b0} - i_{p0}$.

Expressions like $|r_p + Z_{L(m)}|$ indicate that the magnitude only is wanted, *i.e.*, $|r_p + Z_{L(m)}| = \sqrt{(r_p + R_{L(m)})^2 + X_{L(m)}{}^2}$, where $Z_{L(m)} = R_{L(m)} + jX_{L(m)}$.

[1] For the sake of brevity we shall refer to $2\pi f$ as an angular frequency.

A similar analysis shows that a_2 in the third term of Eq. (12-32) is $a_{2(\omega+\omega)}$ and $Z_{L(m\pm n)}$ in Eq. (12-47) becomes $Z_{L(\omega+\omega)}$ and must be evaluated at the angular frequency $(\omega + \omega) = 2\omega$.

To illustrate further the use of Eqs. (12-46) and (12-47) let us assume a power amplifier, transformer-coupled to a resistance load. If properly designed, the transformer may be considered ideal throughout the useful frequency range of the amplifier, introducing a pure resistance load R_L into the plate circuit, whereas at zero frequency the impedance introduced into the plate circuit is essentially zero. With a sine-wave voltage applied to the input, Eq. (12-28), the incremental plate current will be given by Eq. (12-32) with proper modification of a_1 and a_2. Under these conditions a_2 in the first term of Eq. (12-32), assuming μ to be reasonably constant, becomes

$$a_{2(\omega-\omega)} = \frac{\mu^2 r_p r_p'}{2r_p(r_p + R_L)^2} = \frac{\mu^2 r_p'}{2(r_p + R_L)^2} \qquad (12\text{-}48)$$

since $Z_{L(m)} = R_L$, $Z_{L(n)} = R_L$, and $Z_{L(m-n)} = 0$.

The a_2 appearing in the third term is $a_{2(\omega+\omega)}$ and is given by Eq. (12-47) since the impedance at the angular frequency $(m + n) = (\omega + \omega) = 2\omega$ is R_L, not zero as at frequency $(\omega - \omega)$. Thus a_2 in the first term of Eq. (12-32) is appreciably larger than a_2 in the third term, bearing out the statement in the footnote on page 495.

No mention has been made of the third-order term a_3 in Eq. (12-32). Actually it, too, is affected in a similar manner, but no attempt is made here to derive or present its equation. If derived, the equation would be a function of three angular frequencies, m, n, and p, and a_3 would be written $a_{3(m\pm n\pm p)}$. In Eq. (12-32) the a_3 terms are formed from the expansion of $\sin^3 \omega t$ so that all three angular frequencies are equal to ω. Thus a_3 in the term $(3a_3E_{gm}^3/4) \sin \omega t$ is really $a_{3(m+n-p)}$ or $a_{3(\omega+\omega-\omega)}$ while in the term $(-a_3E_{gm}^3/4) \sin 3\omega t$ it is really $a_{3(m+n+p)}$ or $a_{3(\omega+\omega+\omega)}$.

Effect of Applying Two Simultaneous Signals of Different Frequency. Let us now consider the application of these equations when two or more signals of different frequency are applied to the tube simultaneously. Let us assume that e_g (or e_s as it was called in the diode development) is given by

$$e_g = E_{1m} \sin \omega t + E_{2m} \sin qt \qquad (12\text{-}49)$$

Substituting this equation into Eq. (12-5) gives

$$i_{p0} = a_1 E_{1m} \sin \omega t + a_1 E_{2m} \sin qt + a_2 E_{1m}^2 \sin^2 \omega t + 2a_2 E_{1m} E_{2m} \sin \omega t \sin qt + a_2 E_{2m}^2 \sin^2 qt \quad (12\text{-}50)$$

The frequency subscripts have not yet been added to a_1 and a_2 since they cannot be assigned until each term has been reduced to include only a single frequency term. This requires that the equation be expanded through the aid of Eq. (12-30) and the equation for the product of two sines

$$\sin \omega t \sin qt = \tfrac{1}{2}[\cos (\omega - q)t - \cos (\omega + q)t] \quad (12\text{-}51)$$

When this is done, the following equation is obtained in which the frequency subscripts, $m = \omega$ and $n = q$, have been included,

$$i_{p0} = a_{1(\omega)} E_{1m} \sin \omega t + a_{1(q)} E_{2m} \sin qt + a_{2(\omega-\omega)} \frac{E_{1m}^2}{2} - a_{2(\omega+\omega)} \frac{E_{1m}^2}{2} \cos 2\omega t + a_{2(\omega-q)} E_{1m} E_{2m} \cos (\omega - q)t - a_{2(\omega+q)} E_{1m} E_{2m} \cos (\omega + q)t + a_{2(q-q)} \frac{E_{2m}^2}{2} - a_{2(q+q)} \frac{E_{2m}^2}{2} \cos 2qt \quad (12\text{-}52)$$

Proper assignment of the m and n subscripts in this equation may be facilitated by using Eq. (12-51) for the expansion of the $(\text{sine})^2$ terms instead of Eq. (12-30). This may be done by considering $\sin^2 \omega t$, for example, as $(\sin \omega t)$ $(\sin \omega t)$ which, when inserted in Eq. (12-51), gives

$$\sin \omega t \sin \omega t = \tfrac{1}{2}[\cos (\omega - \omega)t - \cos (\omega + \omega)t] = \tfrac{1}{2}(1 - \cos 2\omega t) \quad (12\text{-}53)$$

Comparing this equation with Eq. (12-30) shows that both give the same result, but Eq. (12-53) clearly indicates that the d-c term is formed by the difference between ω and ω; therefore, the frequency subscript must be $(\omega - \omega)$ as shown in Eq. (12-52). A similar analysis may be applied to the other terms containing the frequency subscripts $(\omega + \omega)$, $(q - q)$, and $(q + q)$.

The a_1 and a_2 terms in Eq. (12-52) must be evaluated by means of Eqs. (12-46) and (12-47) to determine the actual magnitude of the various components in the incremental plate current, i_{p0}.

It should be noted that Eq. (12–52) includes components due to the first two terms only of Eq. (12-5), and normally there are additional components due to higher order terms, although these are commonly much smaller than those included in Eq. (12-52) and may usually be neglected. The total plate current, i_b, is found by subtracting i_{p0} of Eq. (12-52) from the quiescent direct current I_{b0}, as indicated by Eq. (12-3).

Determination of r_p', μ_b', and μ_c'. Before evaluating the a_2 coefficients from Eq. (12-47), it is necessary to determine r_p' and, unless μ may be considered constant, μ_b' and μ_c'. A simple procedure is to measure r_p and μ for a number of plate voltages and μ for a number of grid voltages in the vicinity of the direct potentials which are to be applied to the tube. Measurement of the slope of the resulting curves at the desired direct potentials will yield results sufficiently accurate for most purposes.

Example. As an example of the use of the power-series expansion for triode tubes, let us assume a single-stage amplifier with a load impedance consisting of a pure inductance of 2 henrys. Let the tube coefficients be $\mu = 3.5$, $r_p = 3000$ ohms, $r_p' = -300$ ohms/volt, $\mu_b' = 0$, and $\mu_c' = 0$. Let the applied grid voltage be given by Eq. (12-49) in which $\omega = 2\pi 500$ and $q = 2\pi 800$. The various coefficients of Eq. (12-52) may be evaluated with the aid of Eqs. (12-46) and (12-47) in which $Z_L = j2\pi f(2)$, where $2\pi f$ is equal to the subscript indicated in Eq. (12-52). Thus we may write from Eq. (12-46)

$$a_{1(\omega)} = -\frac{3.5}{\sqrt{(3000)^2 + (6280)^2}} = 505 \times 10^{-6} \text{ amp/volt}$$

since $Z_{L(\omega)} = j2\pi 500(2) = j6280$. We may also write

$$a_{1(q)} = -\frac{3.5}{\sqrt{(3000)^2 + (10{,}000)^2}} = 335 \times 10^{-6} \text{ amp/volt}$$

since $Z_{L(q)} = j2\pi 800(2) = j10{,}000$. As a sample of the method of computing the a_2 coefficients from Eq. (12-47) consider the following:

$$a_{2(\omega-\omega)} = \frac{-(3.5)^2 \times 3000 \times 300}{2\sqrt{(3000)^2 + (6280)^2}\sqrt{(3000)^2 + (6280)^2}\sqrt{(3000)^2 + (0)^2}}$$
$$= 38.2 \times 10^{-6} \text{ amp/volt}^2$$

since $Z_{L(\omega-\omega)}$ is the impedance at zero frequency which is zero. Similarly

$$a_{2(\omega+\omega)} = \frac{-(3.5)^2 \times 3000 \times 300}{2\sqrt{(3000)^2 + (6280)^2}\sqrt{(3000)^2 + (6280)^2}\sqrt{(3000)^2 + (12{,}560)^2}}$$
$$= 8.8 \times 10^{-6} \text{ amp/volt}^2$$

$$a_{2(\omega-q)} = \frac{-(3.5)^2 \times 3000 \times 300}{2\sqrt{(3000)^2 + (6280)^2}\sqrt{(3000)^2 + (10{,}000)^2}\sqrt{(3000)^2 + (3760)^2}}$$

$$= 15.8 \times 10^{-6} \text{ amp/volt}^2$$

The remaining a_2 coefficients may be computed in a similar manner.

Diodes. As stated on page 500 the performance of a diode may be determined by using the equations of the triode but with μ considered as being equal to unity. This means that the derivatives of μ with respect to e_b and e_c are zero, and Eq. (12-47) may be used by omitting all but the first term in the numerator. This, of course, gives the same equation as Eq. (12-17) with $\mu = 1$ but with the denominator expanded to take account of the variation of impedance with frequency, as was done for the triode in the preceding sections.

Pentode and Beam Tubes. The performance of pentode and beam tubes, with load impedances which are a function of frequency, may be computed with the aid of Eqs. (12-46) and (12-47) in exactly the same manner as for the triode. However Z_L is nearly always much smaller than r_p, and the approximate Eqs. (12-10), (12-24), and (12-25) may be used. These are entirely independent of the load impedance, and no problem is involved in applying them to a circuit where the load impedance varies with the frequency.

Components Produced by Distortion. Equation (12-52) contains eight terms, each of which was produced by one of the impressed components—$E_{1m} \sin \omega t$ and $E_{2m} \sin qt$—or by a combination of the two. The results of square-law distortion, with two impressed components, may therefore be generalized with respect to the frequencies of the various components in the output by saying that terms will appear having frequencies equal to (1) those of the original components, (2) the sum of the frequencies of each pair of original components, (3) the difference between the frequencies of each pair of original components. The pairs used to find the sum and difference terms must include each original component paired with itself as well as with the other component. To illustrate, the original components from which Eq. (12-52) was derived were of angular frequency ω and q. Equation (12-52) includes the following terms: (1) ω, q; (2) $\omega + \omega = 2\omega$, $\omega + q$, $q + q = 2q$; (3) $\omega - \omega = 0$, $\omega - q$, $q - q = 0$. The zero frequency terms are of course the two direct components.

It may be shown that the foregoing analysis applies to distortion when any number of components are present in the input signal. Thus if three components are impressed having angular frequencies of ω, q, and p, the angular frequencies of the resulting terms will be (1) ω, q, p; (2) $\omega + \omega = 2\omega, \omega + q, \omega + p, q + q = 2q, q + p, p + p = 2p$; (3) $\omega - \omega = 0, \omega - q, \omega - p, q - q = 0, q - p, p - p = 0$

The foregoing concepts are valuable in predicting, without going through the expansion that produced Eq. (12-52), the exact frequency of each component that may be expected as the result of amplitude distortion. The analysis may even be expanded to include higher order terms. For the third order, additional terms must include frequencies resulting from all possible combinations of any three components in the input such as $\omega + \omega + \omega = 3\omega, \omega + \omega - \omega = \omega, \omega + \omega + p = 2\omega + p, \omega + \omega - p = 2\omega - p, \omega + p + q, \omega + p - q, \omega - p + q, p + p + q = 2p + q$, etc

Applications of the Power-series Analysis. The power-series analysis may be used to analyze amplitude distortion in amplifiers of all types, although when resistance loads are used the graphical method of Chap 9 is simpler If a reactive load is used, the graphical method is not practical, since the load line becomes an ellipse, and the power-series approach may be helpful. The power-series analysis is especially helpful in treating square-law modulation and demodulation (see Chaps 13 and 14) and in solving special problems of distortion. In particular, it is an aid in determining the frequencies of the new components added by amplitude distortion of a known wave.

Problems

12-1. Evaluate a_1 and a_2 in Eq (12-5) for a tube with characteristics as in Fig 9-15, $\mu = 3.5$ and $R_L = 3000$ ohms Assume the operating point to be at $E_{bo} = 275$ volts, $E_c = -56$.

HINT Measure r_p and r_p' from the curve by graphical methods

12-2. For a given triode tube operating at certain direct potentials $\mu = 15$, $r_p = 50{,}000$ ohms, $r_p' = -2000$ ohms/volt Compute the fundamental and second harmonic voltages appearing across an output resistance of 50,000 ohms with an impressed voltage on the grid of $E_g = 0.141 \sin 2\pi 10^7 t$ (Assume μ to be constant)

12-3. Repeat Prob 12-2 for a load consisting of a parallel-resonant circuit in which $C = 31.8\ \mu\mu f$ and L is of such size as to provide resonance at the frequency of the impressed signal. L has a Q of 100.

12-4. Repeat Prob. 12-2 for a load consisting of a resistor and condenser in parallel, the resistor having a resistance of 70,700 ohms and the condenser being of such size that its reactance is 70,700 ohms at the frequency of the impressed signal.

12-5. Repeat Prob. 12-2 for a load consisting of a resistor and coil in series, the resistor having a resistance of 35,350 ohms and the coil having a reactance of 35,350 ohms at the frequency of the impressed signal.

CHAPTER 13

MODULATORS

One of the most valuable applications of the vacuum tube is that of modulating an alternating current. Modulation makes possible such developments as radiotelephony; carrier-current telephony over land wires, both telephone and power; remote metering over power wires; and many other similar applications.

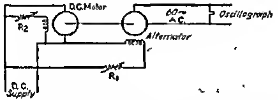


FIG. 13·1 A 60-cycle alternator with provision for producing amplitude modulation by varying R_1, or frequency modulation by varying R_2

Modulation. Modulation is defined in the 1938 Standards Report of the Institute of Radio Engineers as "the process of producing a wave some characteristic of which varies as a function of the instantaneous value of another wave, called the modulating wave."

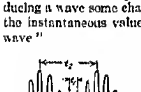


FIG. 13-2 Amplitude-modulated wave. Sinusoidal modulation.

A very simple illustration of modulation is shown in Fig. 13-1. If the field rheostat of the 60-cycle alternator is varied periodically at, say, 10 cycles/sec, its output will be a 60-cycle wave varying in amplitude at the lower frequency of 10 cycles/sec. An oscillogram of this output is sketched in Fig. 13-2, where t_1 represents one cycle of the 60-cycle, or carrier, wave and t_2 represents one cycle of the 10-cycle, or modulating, wave. In this type of modulation the amplitude of the 60-cycle output is varied, while the period and phase remain

constant; it is therefore known as *amplitude modulation*. This is the type most commonly used in practice at the present time.

Suppose that the output voltage of the alternator, referred to in the preceding paragraph, is maintained constant by a voltage regulator and that the machine is driven by a d-c shunt motor, the field rheostat of which is varied at a rate of 10 cycles/sec. The resulting wave will appear as in Fig. 13-3, where the amplitude of the resulting wave remains constant but the frequency varies at a 10-cycle rate[1] between definite upper and lower limits, around 60 cycles as an average. This type of modulation is known as *frequency modulation*.

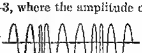


FIG. 13-3. Frequency-modulated wave. Sinusoidal modulation.

Finally, suppose that the alternator is connected in parallel with a very large 60-cycle system, the connection being made through large series reactances, as in Fig. 13-4. Again let the field current of the driving motor be varied periodically. The frequency

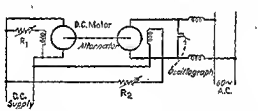


FIG. 13-4. A 60-cycle alternator with provision for phase modulation by varying R_1.

of the alternator must remain the same as that of the system to which it is tied, but its phase will vary periodically with reference to that of the associated system. This type of modulation is known as *phase modulation*.

As a matter of fact, the resultant wave produced by phase modulation is very similar to that produced by frequency modulation,[2] since any change in either frequency or phase must necessarily produce at least a momentary change in the other. For this reason phase modulation is sometimes treated as frequency

[1] Assuming, for the sake of the example, that the mass of the system is sufficiently low to enable the machine to respond to such rapid variations.

[2] Hans Roder, Amplitude, Phase, and Frequency Modulation, *Proc. IRE*, **19**, p. 2145, December, 1931.

modulation, although this is not a strictly correct point of view as is shown later (page 538).

In the foregoing examples of modulation the wave produced by the alternator when the rheostats R_1 and R_2 are at rest in their mid-positions is known as the *carrier wave*, because in radio transmission it is of sufficiently high frequency to cause radiation of energy through space and thus "carries" the lower voice frequencies (or whatever may be used to modulate the carrier wave) from the transmitter to the receiver. The wave represented by the periodic variation of either R_1 or R_2 is known as the *modulating wave* and is normally of a very much lower frequency than that of the carrier wave.

Uses of Modulation. The most outstanding applications of modulation are in the field of communication, where it is used in the transmission of voice, music, pictures, and other forms of information by means of radio and wire. In the application to wire, several individual communications may be carried on simultaneously over the same pair of wires by transmitting currents of high audio or supersonic frequencies modulated by the information-carrying frequencies. Such a system is generally known as *carrier current* or simply *carrier*. It may also be applied to the transmission of voice over 60-cycle power lines so that the same set of wires will transmit power and provide communication between power plants and substations. In the application to radio, the information-carrying frequencies are modulated or superimposed on higher radio frequencies which are capable of establishing radiating fields to carry the information over great distances without the aid of interconnecting wires. The modulating frequencies are in themselves too low to produce a radiating field, so that only through the medium of modulation is radio communication possible. This communication may take the form of telegraphy, telephony, facsimile, or television, all of which utilize the same fundamental principles of modulation and demodulation. Similar principles apply to radio aids to navigation such as radar, loran, and direction finders, although not all these applications require modulation.

1. AMPLITUDE MODULATION

The Standards Report of IRE defines an amplitude-modulated wave as "one whose envelope contains a component similar to the wave form of the signal to be transmitted." If the wave shape

of the envelope is exactly like that of the signal, modulation is said to be *linear* and no distortion is present.

We may write the equation of an unmodulated carrier wave as

$$e_0 = E_{0m} \sin \omega t \qquad (13\text{-}1)$$

which is a wave such as shown in Fig. 13-5. The crest voltage E_{0m} and the time of one cycle $2\pi/\omega$ are indicated. The dotted line shows the shape of the envelope, which is a straight line lying above the zero line by an amount E_{0m}.

The definition of a linearly modulated wave, given in a preceding paragraph, indicates that the wave of Fig. 13-5, when modulated, should appear as in Fig. 13-6, where m is merely a multiplying factor which is a function of the amplitude of the

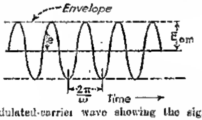


FIG. 13-5. Unmodulated-carrier wave showing the significance of the terminology.

modulating wave. From inspection of this figure we may write the equation of an a-m wave as[1]

$$e = (E_{0m} + mE_{0m} \sin qt) \sin \omega t \qquad (13\text{-}2)$$

where the portion of the equation inside the parentheses is the equation of the envelope and merely replaces E_{0m} in Eq. (13-1). Taking E_{0m} out of the parentheses gives

$$e = E_{0m}(1 + m \sin qt) \sin \omega t \qquad (13\text{-}3)$$

The term m is known as the *modulation factor* and is defined in the 1938 Standards Report of IRE, for amplitude modulation, as "the ratio of half the difference between the maximum and minimum amplitudes to the average amplitude."[2]

[1] The abbreviation "a-m" is commonly used for "amplitude-modulated."

[2] A footnote to this definition is as follows: "In linear modulation the average amplitude of the envelope is equal to the amplitude of the unmodulated wave, provided there is no zero-frequency component in the modulat-

Side Bands. If the multiplication process indicated by Eq. (13-3) is carried out, the result is

$$e = E_{0m} \sin \omega t + mE_{0m} \sin \omega t \sin qt \qquad (13\text{-}4)$$

This equation may be expanded further by using the trigonometric expression for the product of two sines to give

$$e = E_{0m} \sin \omega t - \frac{mE_{0m}}{2} \cos (\omega + q)t + \frac{mE_{0m}}{2} \cos (\omega - q)t \quad (13\text{-}5)^*$$

Examination of this equation shows the presence of three separate and distinct components of different frequencies: (1) the original carrier frequency at its original amplitude, (2) a component having a frequency equal to the sum of the carrier and modulating frequencies, and (3) a component having a frequency equal to the difference between the carrier and modulating frequencies.* The second term is known as the *upper side frequency;* the third is known as the *lower side frequency.* The modulating wave rarely consists of a single frequency as assumed in the preceding analysis but more commonly consists of a number of components having different frequencies, so that the second and third terms in Eq. (13-5) consist of a band of frequencies lying between the carrier frequency plus or minus the lowest modulating frequency and the carrier frequency plus or minus the highest modulating frequency. These are known as *side bands.*

FIG. 13-6. Amplitude modulated wave.

The presence of side frequencies may be detected experimentally, if the ratio of carrier to modulating frequency is not too large, by means of a series-tuned circuit and suitable current indicator. As the test circuit is brought into resonance in turn with each side frequency and the carrier, very definite increases in current may be observed, corresponding to each of the components of Eq. (13-5).

ing signal wave (as in telephony). For modulating signal waves having unequal positive and negative peaks, positive and negative modulation factors may be defined as the ratios of the maximum departures (positive and negative) of the envelope from its average value, to its average value."

Power in a Modulated Wave. The average power in any circuit is given by

$$P = I^2R \tag{13-6}$$

where I is the effective value of the current in the circuit and R the equivalent resistance. The instantaneous current produced in a circuit of resistance R by the emf of Eq. (13-5) is evidently

$$i = I_{0m} \sin \omega t - \frac{mI_{0m}}{2} \cos (\omega + q)t + \frac{mI_{0m}}{2} \cos (\omega - q)t \quad (13\text{-}7)^*$$

where $I_{0m} = E_{0m}/R$.

The rms value of this current may be found by taking the square root of the sum of the squares of the rms values of its various components.[1] The I^2 of Eq. (13-6) will, therefore, in the case of a modulated wave, be

$$I^2 = I_0^2 + \frac{m^2I_0^2}{4} + \frac{m^2I_0^2}{4} \tag{13-8}$$

and the power of the modulated wave will, from Eq. (13-6), be

$$P = I_0^2R + \frac{m^2I_0^2R}{2} \quad (13\text{-}9)^*$$

If m is zero, *i.e.*, if no modulation is present, Eq. (13-9) will give the power in the carrier wave alone

$$P = I_0^2R \tag{13-10}$$

Therefore, the power in an a-m wave is greater than that of the unmodulated carrier by the power contained in the side bands, or $P_{out} = P_{carrier} + P_{side\ bands}$ and the side-band power is

$$P_{sb} = \frac{m^2I_0^2R}{2} = \frac{m^2P_c}{2} \quad (13\text{-}11)^*$$

where P_{sb} is the side-band power and P_c is the carrier power. The power in the side bands is proportional to the square of the modulation factor and has a maximum possible value of one-half the carrier energy.[2] Since only that portion of the power contained in the side bands carries any information, it is evident that a high modulation factor is highly desirable.

[1] See any text on a-c theory.

[2] m cannot exceed 1 if the modulation is to be linear.

Linear Modulators. A linear modulator is one that produces a modulated wave conforming to Eq. (13-5) or (13-7). It was shown in the development of Eq. (13-3) that modulation is linear when the crest value of the r-f output voltage is proportional to $(1 + m \sin qt)$. The first term of this expression is a constant; the second is proportional to the modulating signal. Therefore, linear modulation is produced by a vacuum-tube amplifier when the amplitude of the r-f output voltage of the amplifier is directly proportional to the sum of a direct (or constant) voltage and an alternating modulating voltage applied to an electrode of the tube.

Linear Plate Modulation. Triodes. Probably the most common method of producing linear modulation of vacuum tubes is that of plate modulation. As its name implies, this method makes use of the dependence of the output of a tube upon its plate voltage. If a curve of r-f output voltage vs. direct plate voltage is taken on a class C triode amplifier, the result is substantially a straight line (Fig. 13-7). Therefore, if the plate voltage is varied periodically about some average value, the output voltage and current will also vary periodically, producing a wave like that of Fig. 13-2. A simple circuit for producing plate modulation is shown in Fig. 13-8, where a source of carrier voltage is applied to the grid circuit of a class C amplifier with a source of modulating voltage $E_{am} \sin qt$ in series with the d-c supply to the plate. Evidently the average power delivered by the plate power supplies—both d-c and a-c—is

FIG. 13-7. Output voltage vs. direct plate voltage for a class C amplifier.

$$P_s = \frac{1}{2\pi} \int_0^{2\pi} \frac{(E_b + E_{am} \sin qt)^2}{R_b}\, d(qt) = \frac{E_b^2}{R_b} + \frac{E_{am}^2}{2R_b} \qquad (13\text{-}12)$$

where R_b = equivalent d-c plate resistance of amplifier tube, which is essentially constant[1]

E_{am} = crest value of modulating voltage

q = 2π times modulating frequency

[1] R_b may be used with the alternating modulating voltage rather than r_p, since the curve of average plate current (averaged over an r-f cycle) vs. applied plate voltage must be straight for linear modulation, and the a-c resistance presented to the signal voltage is therefore essentially equal to the d-c resistance R_b.

Evidently the first term in the right-hand side of Eq. (13-12) is the power supplied by the battery, whereas the second term is that supplied by the source of modulating voltage.

The output of the tube P_0, given by Eq. (13-9), may be related to the input as follows:

$$\eta P_i = P_0 \tag{13-13}$$

where η = efficiency of the amplifier.

Insertion of the values of P_0 and P_i [from Eqs. (13-9) and (13-12), respectively] into Eq. (13-13) gives

$$\eta\left(\frac{E_b^2}{R_b} + \frac{E_{am}^2}{2R_b}\right) = I_0^2R + \frac{m^2I_0^2R}{2} \tag{13-14}$$

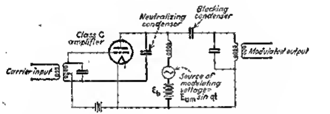


FIG. 13-8. Circuit illustrating the method of applying plate modulation.

R being the resistance of the load circuit. If the signal voltage is now removed, Eq. (13-14) reduces to[1]

$$\eta\frac{E_b^2}{R_b} = I_0^2R \tag{13-15}$$

If Eq. (13-15) is subtracted from Eq. (13-14), the result is

$$\eta\frac{E_{am}^2}{2R_b} = \frac{m^2I_0^2R}{2} \tag{13-16}$$

The conclusion to be drawn from Eqs. (13-15) and (13-16) is evidently that *the carrier power is supplied by the source of direct plate voltage whereas the side-band power is supplied by the source of*

[1] This, of course, assumes that the efficiency remains unchanged when the modulating voltage is removed. That this is a reasonable assumption may be seen from Fig 13-7, since the efficiency is proportional to the ratio of output to input power and therefore essentially proportional to (output voltage)$^2/E_b^2$.

modulating voltage. Both sources also supply the losses of the class C amplifier in the ratio of side band to carrier power $m^2/2$. It is further evident that for 100 per cent modulation the second term on the right-hand side of Eq. (13-14) is half the first; therefore, the second term on the left-hand side of the same equation must also be half the first or, for 100 per cent modulation ($m = 1$),

$$E_{sm} = E_b \tag{13-17}$$

From the foregoing analysis it may be concluded that the source of modulating voltage must have a power capacity of 50 per cent of the unmodulated input power to the modulated amplifier and that the losses in the modulated amplifier will be increased by 50 per cent if 100 per cent modulation is applied. It may also be seen that the peak value of plate voltage will be *double* the direct voltage under these conditions and the peak power will be *four times the carrier power.* Evidently such considerations must be taken into account in designing the class C amplifier that is to be modulated, or overloading will result.[1]

Sources of Modulating Voltage for Plate-modulated Amplifiers. The source of signal or modulating voltage shown in Fig. 13-8 may be any generator or other device capable of delivering the voltage and power required, usually a vacuum-tube amplifier. The original source of the signal voltage is commonly a microphone or other low-power device, so that a reasonable amount of amplification is generally required. The modulator may, therefore, be preceded by any number of stages of amplification.

The Heising system of modulation, utilizing a vacuum-tube amplifier as modulator, is shown in Fig. 13-9. The modulator tube is biased to operate as a class A power amplifier in order to reproduce accurately all the original frequency components without distortion. The transformer T should have a turns ratio such as to reflect the proper load resistance for maximum output at low distortion, as in the class A power amplifier study of Chap. 9 (starting on page 319).[2] The correct bias should also be deter-

[1] An excellent method of designing a class C modulated amplifier is given by Wagener, similar to the design of an unmodulated class C amplifier given in Chap. 10 (starting on p. 405). The reader is referred to Wagener's paper for further details.

[2] The modulated class C amplifier presents a load resistance across the secondary of the modulation transformer of essentially $R_c = E_b/I_b$.

mined in the manner described in that study. The inductance L_1 is a r-f choke designed to prevent the flow of currents of carrier frequency through the modulator circuits.

✓The modulator tube in the circuit of Fig. 13-9 must be of considerably larger capacity than the class C amplifier that it modulates. A simple method of demonstrating this fact is to work out an example. Assume that the class C amplifier delivers a carrier output of 100 watts and that its efficiency is 75 per cent. When unmodulated, the input to the class C amplifier is 133 watts and its plate loss 33 watts. With sufficient signal applied to the grid of the modulator to produce 100 per cent modulation the output of the class C amplifier is increased by 50 per cent, making a total output for both carrier and side bands of 150 watts. The input to the class C amplifier is now 200 watts, and its plate loss is 50

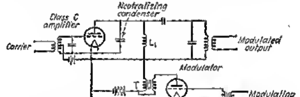


FIG. 13-9. Heising-modulator circuit using a modulation transformer.

watts, since the efficiency is still 75 per cent. Of the 200 watts input, 133 watts is supplied directly by the d-c source of power, whereas the remaining 67 watts represents output from the class A modulator. If the efficiency of the class A modulator is 25 per cent, its input is 267 watts and its plate loss 200 watts. Thus the modulator tube must be capable of dissipating 200 watts on its plate, but the class C tube need dissipate only 50 watts. As a matter of fact, the plate-dissipating capacity of the modulator should be equal to its input power of 267 watts, since it was shown in Chap. 9 that the input to a class A amplifier is constant regardless of its output. Thus if the signal voltage applied to the grid of the modulator is removed (as during a lull in the transmission of voice or music), the entire 267 watts must be dissipated by the plate of the tube. For the efficiencies assumed, then, the modulator tube must be capable of dissipating over *five times* as much power as is required of the class C amplifier.

The foregoing example is summarized in Table 13-1.

The modulation transformer of Fig. 13-9 must be capable of passing the comparatively large direct current drawn by the two tubes without overheating and without saturation of the iron core.[1] It must also be insulated for comparatively high voltages, since, for 100 per cent modulation, the crest value of alternating voltage appearing across its output terminals must equal the direct plate voltage of the class C amplifier, whereas the crest value of voltage to ground is the sum of the direct and alternating voltages or twice the direct voltage.

Plate Modulation with Class B Modulators. The large power capacity required of the class A modulator of the preceding section

TABLE 13-1 EXAMPLE OF PLATE-DISSIPATION REQUIREMENT OF THE MODULATOR AND CLASS C AMPLIFIER IN THE CIRCUIT OF FIG. 13-9

	Class C amplifier Efficiency = 75 per cent			Class A modulator Efficiency = 25 per cent		
	Power input to plate	Loss on plate	Power output	Power input	Loss on plate	Power output
No signal voltage on grid of modulator . .	133	33	100	267	267	0
Signal voltage on grid of modulator sufficient to produce 100 per cent modulation.... .	200	50	150	267	200	67

was due to its low efficiency. It, therefore, seems logical to use class B (or class AB) push-pull amplifiers as modulators, Fig. 13-10, whenever appreciable power is to be handled. Not only is the efficiency of the class B amplifier higher than that of the class A but the zero-signal power input is much less than the input for maximum signal. Thus if, in the example of the preceding section, a class B modulator is used with an efficiency of 65 per cent at maximum signal voltage, the power output for 100 per cent

[1] The transformer should be so connected that the d-c magnetizations set up by the plate currents of the two tubes are in phase opposition, thus materially reducing saturation problems.

modulation must still be 67 watts but the input is only 102 watts and the plate loss 35 watts. Since the zero-signal input is much less than 35 watts, this modulator need be designed to dissipate a maximum of only 35 watts instead of the 267 watts required of the class A modulator. Thus much smaller tubes may be used.

Class AB modulators are more easily adjusted to give low distortion than are class B and are, therefore, commonly used in medium-power installations where their somewhat lower efficiency is not so much of a disadvantage as in high-power units. The circuit is, of course, the same as that of Fig. 13-10, but with a different bias adjustment than for a class B amplifier. For that matter, the same circuit may be used for a push-pull class A amplifier if desired.

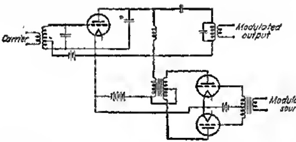


FIG. 13-10. Circuit of a class B push-pull amplifier used as a Heising modulator.

Distortion in Plate-modulated Class C Amplifiers. It is evident that the amount of distortion introduced by plate modulation of class C amplifiers is a function of the linearity of the curve shown in Fig. 13-7, which is, in turn, a function of the linearity of the characteristic curves of the tube. The curve of Fig. 13-7 always departs somewhat from a straight-line relationship, especially at the lower plate voltages, introducing some distortion.

This lack of linearity may be largely compensated for by causing the bias of the class C amplifier to vary slightly with the modulating voltage. This may be done by supplying a portion of the bias from a fixed source and the remainder from a grid-leak bias, *i.e.*, by inserting a grid-leak and by-pass condenser in series with the grid of the class C amplifier. As the modulating voltage increases toward its positive crest, the total plate voltage applied to the

class C tube increases, causing a decrease in grid current (see Fig. 3-34, page 69) and, therefore, a decrease in bias. On the negative half cycle of the modulating voltage the grid current, and therefore the bias, increases.

The need for a change in bias may be seen by first considering the performance of the class C amplifier without modulation. Let the direct plate voltage and crest r-f plate voltage, under these conditions, be E_b and E_{pm}, respectively. We may therefore write $E_b = E_{pm} + e_{b\ min}$. If the amplifier has been designed for high efficiency under carrier conditions, we may also write $e_{b\ min} = e_{c\ max}$.

Consider next the performance at the crest of the modulation cycle. The effective direct voltage (normal direct plus the crest value of the a-f modulating voltage) is now $2E_b$, and the crest r-f voltage, for linear operation, should be $2E_{pm}$. The minimum plate voltage is the difference between these two and is therefore twice the no-modulation $e_{b\ min}$. But $e_{c\ max}$ *is unchanged from the no-modulation condition and is now only half as large as* $e_{b\ min}$, *whereas it should be equal to it.* A decrease in bias will increase $e_{c\ max}$ and so restore the tube to normal operation. It is not necessary to decrease the bias to a point where $e_{c\ max}$ again equals $e_{b\ min}$, to secure reasonably linear operation, but some decrease is desirable.[1]

Plate Modulation of Tetrode and Pentode Tubes. It is not possible to produce satisfactory modulation of tetrode and pentode tubes by merely inserting a modulating voltage in series with the plate. The plate current in these tubes is nearly independent of the plate voltage, whereas both plate current and voltage must vary directly with the modulating voltage if the ratio of power output to power input is to be constant, as in the class C plate-modulated *triode* amplifier. However, plate current varies with the screen-grid voltage of a tetrode or pentode in the same manner as with the plate voltage of a triode;[2] therefore, if the modulating voltage is applied to the screen as well as to the plate (Fig. 13-11), tetrodes and pentodes may be modulated in the same manner as triodes.

[1] For further details, see W. G. Wagener, Simplified Methods for Computing Performance of Transmitting Tubes, *Proc. IRE*, 25, p. 47, January, 1937.

[2] This is true of tetrodes as long as the plate voltage exceeds that of the screen grid.

Another method of modulating tetrodes and pentodes is to apply the modulating voltage to the screen grid alone, but the distortion produced is greater than with the method of Fig. 13-11.[1] In general, best results are obtained by using triodes in modulated class C amplifiers.

Grid Modulation.[2] Class C amplifiers may also be modulated by applying the modulating signal to the grid of the tube rather than to the plate (Fig. 13-12). Evidently this permits the use of a lower power modulator tube (not shown) than when the modu-

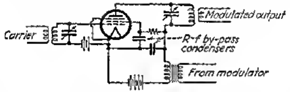


FIG. 13-11. Circuit for plate-modulating a pentode tube by applying the modulation voltage to both the plate and the screen grid.

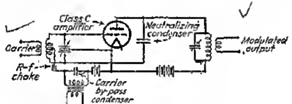


FIG. 13-12. Circuit of a grid-modulated class C amplifier.

lating voltage is applied to the plate of the amplifier, since the power demand in the grid circuit of a class C amplifier is far less than that in the plate. The distortion produced with this type of modulation tends to be a little higher than with the Heising method, although it may be largely eliminated by the introduction of feedback.

[1] For a study of the various methods of modulating screen grid tubes, see H. A. Robinson, An Experimental Study of the Tetrode as a Modulated Radio-frequency Amplifier, *Proc. IRE*, **20**, p. 131, January, 1932.

[2] See also F. E. Terman and R. R. Buss, *Proc. IRE*, **29**, p. 104, March 1941.

Figure 13-13 illustrates the process of grid modulation. When no modulating voltage is applied to the grid of the class C amplifier, the grid voltage and plate current are as drawn between points a and b, the tube operating as a regular class C amplifier. When modulating voltage is applied, the effect is as though the grid bias were varied around its normal value as shown by the modulating wave between b and c. At the positive crest of modulation cycle (at d) the effective bias (d-c bias plus modulation voltage) is approximately equal to the cutoff voltage and the class C amplifier passes current for about 180 deg, of the carrier

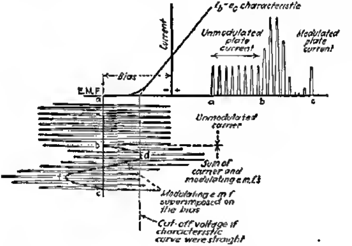


FIG 13-13 Illustrating the manner in which modulation is attained in the circuit of Fig 13-12

cycle producing maximum power output. At the negative crest of the modulation cycle (at f) the effective bias is such that the r-f driving voltage is just able to swing the grid to cutoff at its positive crest. Thus the output of the class C amplifier varies periodically between a maximum and zero, giving a 100 per cent modulated wave the linearity of which depends on the shape of the i_b-e_c dynamic characteristic curve.

It is evident from the foregoing discussion and from Fig. 13-13 that the r-f plate voltage of the amplifier (which has the appear-

ance of Fig. 13-6) is a maximum at the positive crest of the modulating voltage and zero at the negative crest, for 100 per cent linear modulation. The amplifier should therefore be designed to operate at high efficiency at d, the point in the modulating cycle at which the output is a maximum; *i.e.*, the minimum plate voltage $e_{b\ min}$ should be approximately equal to the maximum positive grid voltage $e_{c\ max}$ at point d. As the modulating voltage passes through its half cycle from d to f, the r-f output voltage decreases and $e_{b\ min}$ increases, approaching E_b at point f. Therefore, throughout a large part of the modulating cycle, the instantaneous plate voltage is quite high during the time of plate-current flow, resulting in increased plate loss. The efficiency of the amplifier is therefore only about half that of a properly designed *unmodulated* class C amplifier, and the expected output must be correspondingly reduced.[1]

Distortion in Grid-modulated Amplifiers. Driving the grid positive on the modulation peaks as in Fig. 13-13 is necessary if maximum output from the tube is to be realized. On the other

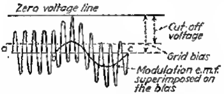


FIG. 13-14. Grid excitation and bias for a class C modulated amplifier in which the grid is not driven positive.

hand, the grid current drawn when the grid swings positive may cause distortion due to impedance drop in the driving source (see page 318). Where maximum quality of transmission is desired, the grid must not be allowed to swing positive, and the grid voltage follows the curve of Fig. 13-14. Here the positive crests of the r-f excitation voltage swing the grid potential just to zero on the positive half of the modulation cycle and just to cutoff on the negative half. The output and efficiency are less than when the grid is driven positive, but this is the price which must be paid for high quality.

[1] These conditions are evidently almost identical with those existing in a class B linear amplifier (p. 432).

Distortion due to nonlinearity of the tube characteristic may be reduced by increasing the load resistance, although this also decreases the power-output capacity of the amplifier. Increasing the load resistance in the plate circuit of a triode amplifier improves the linearity of the dynamic characteristic curve of Fig 13-13, as was demonstrated in Chap 9 (page 327). Distortion of this type may also be reduced by causing the grid bias to vary as a function of the modulating voltage in the same manner as suggested for the plate-modulated amplifier on page 519. Unfortunately this requires the presence of grid current and thus tends to introduce distortion due to impedance drop in the input circuit, unless this impedance is kept sufficiently low by proper design of the circuit (see the discussion on class B amplifiers, page 351).

Tetrodes and pentodes may also be modulated by the grid-modulation method. Their performance is similar to that of the triode, since the control grid in these multigrid tubes operates in a manner similar to that of a triode

Suppressor-grid Modulation of Pentodes. Pentodes may also be modulated by applying a suitable negative bias and the modulation voltage to their suppressor grids (Fig. 13-15). The operation of this circuit is very similar to that of the grid-modulated amplifier of the preceding section. Any change in the suppressor-grid potential varies the flow of current to the plate in a relationship that, if a reasonably high plate-load impedance is used, is quite linear. The negative bias applied to the suppressor grid should be sufficient to reduce the r-f output voltage to half the amplitude obtained with zero voltage on the suppressor, and the crest value of the modulating voltage should be equal to the bias so as not to drive the grid positive. If the suppressor grid is never driven positive, it draws no current and the modulating power required is very low.[1]

A curve of r-f output voltage vs. suppressor-grid potential is shown in Fig. 13-16 for a tube designed especially for suppressor-grid modulation. The proper bias E_{c3} and the maximum permissible crest value of the modulating voltage E_{sm} are indicated.

The plate efficiency of a suppressor-grid-modulated pentode amplifier is about the same as that of a control-grid-modulated amplifier, or about half that of a properly designed, unmodulated class C amplifier. This is because $e_{b\,min}$ can be made equal to

[1] See also C. B. Green, Suppressor-grid Modulation, *Bell. Lab. Record*, 17, p 41, October, 1938.

$e_{c\ \max}$ only at the positive crest of the modulation cycle, just as with control-grid modulation. The *over-all* efficiency is lower with suppressor-grid modulation, however, since the total space current remains virtually unchanged throughout the modulation cycle, whereas it varies with the modulation voltage during control-grid

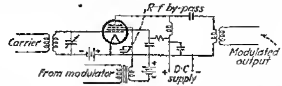


FIG. 13·15. Circuit of a class C pentode amplifier with suppressor-grid modulation.

modulation. Furthermore, since the total space current is essentially constant, the screen current must rise as the plate current drops during the negative half cycle of the modulating voltage, thus producing higher screen-grid losses. This restricts the permissible output of the suppressor-grid modulator.

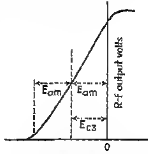


FIG. 13·16. Characteristic curve for the circuit of Fig. 13·15 indicating the degree of linearity which may be expected.

High-efficiency Grid-modulated Class C Amplifier. The efficiency of a grid-modulated or suppressor-modulated class C amplifier may be increased in a manner similar to that used in the Doherty high-efficiency amplifier (page 432). The circuits are basically similar, but the grid excitation, instead of being a modulated wave, is the sum of the carrier and modulating voltages, as in Figs. 13-12 and 13-13.[1]

Use of Cathode-ray Tube in Checking Modulation. One of the simplest methods of checking modulation, both for magnitude and for quality, is through the use of the cathode-ray oscilloscope. The most direct method would be to apply the modulated wave to the vertical deflection plates and a linear sweep circuit (see page 304) to the other pair, producing a picture similar to Fig. 13-6.

[1] See F. E. Terman and John R. Woodyard, A High-efficiency Grid-modulated Amplifier, *Proc. IRE*, **26**, p. 929, August, 1938.

Percentage modulation may be determined with some degree of accuracy,[1] but distortion is not readily observable in this manner. If, instead, the linear sweep circuit is replaced by the a-f modulating voltage, a figure will appear on the screen of the cathode-ray tube similar to that of Fig. 13-17*a* for 100 per cent linear modulation. If the modulation is less than 100 per cent, the figure will appear as in Fig. 13-17*b*; if over 100 per cent, it will appear as in 13-17*c*, with a tail or line extending out to the left of the figure along the X axis, since the output current falls to zero during a portion of the cycle. Nonlinearity will cause a change in the shape of the figure on the cathode-ray tube, as in Fig. 13-17*d*, which evi-

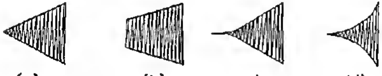


FIG. 13-17. Oscilloscope modulation patterns. (*a*) 100 per cent modulation, no distortion; (*b*) 50 per cent modulation; (*c*) over-modulation; (*d*) nonlinear modulation.

dently indicates a badly curved dynamic characteristic, owing perhaps to too low a load resistance.

Phase shift in the circuit, between the point at which the modulating source is tapped for the horizontal deflection and the point at which the modulation is sampled, will produce an apparent depth to the figure on the screen, the trace made in sweeping from left to right not coinciding with that made in going from right to left. For best results such phase shift should be corrected by introducing a compensating phase shift of opposite sign in the circuit supplying the horizontal plates of the oscilloscope.

Square-law Modulation. Any nonlinear circuit element may produce modulation in an electric circuit. Thus modulation may be produced in a r-f amplifier due to the nonlinearity of its plate-current–grid-voltage characteristic curve. The modulating voltage may be due to inadequate filtering in the rectifier supplying the direct current, or it may be the signal of an undesired radio station which has not yet been fully rejected by the tuned circuits,

[1] See Standards Report of IRE for more accurate methods of determining the percentage of modulation.

as in a radio receiver. This latter phenomenon is more fully discussed in a later section on cross modulation.

Some idea of the process by which modulation of this type takes place may be gathered from Fig. 13-18, where the *i-e* characteristic curve may represent any circuit element in which the current does not vary linearly with the voltage. If the element is a vacuum tube, the curve may represent the current-voltage curve of a diode or the plate-current–grid-voltage curve of a triode or pentode,

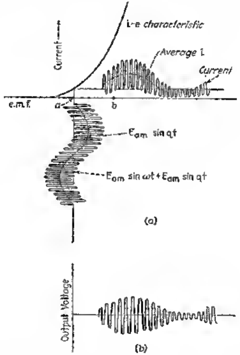


FIG. 13-18. Square-law modulation due to the curvature of the characteristic curve of a vacuum tube.

zero plate voltage being represented by point *a* for the former and zero grid voltage by some such point as *b* for the latter.[1] A suitable direct potential (usually zero for the diode, negative for the grid of the triode or pentode) is applied to the tube, and r-f carrier and a-f modulating voltages of frequency $\omega/2\pi$ and $q/2\pi$,

[1] See first footnote on p. 575.

respectively, are superimposed on the direct component. As may be seen from the figure, the instantaneous ratio of alternating current to impressed alternating voltage varies with the portion of the characteristic curve that is being utilized at any given instant of time, thus producing an output current with a wave shape as shown If this current is passed through a load circuit that presents zero impedance to the modulating and direct components, the voltage across this load will contain only r-f components and will appear as in Fig. 13-18*b*. This voltage is the same general shape as that of Fig 13-2 or 13-6 and is, therefore, an a-m wave.

The van der Bijl modulator of Fig. 13-19 operates on the principle indicated in Fig 13-18 It is essentially a class A amplifier since the plate current flows throughout the cycle (see definition of class A amplifiers on page 602) but is operated on the lower curved portion of the plate-current curve. It has few applications except in the balanced modulator for carrier suppression (see page 533) but many r-f amplifiers act as van der Bijl modulators,

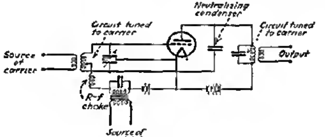


Fig 13-19 Circuit of a square-law triode (van der Bijl) modulator.

and produce undesirable reactions such as cross modulation (see page 531).

The signal levels applied to modulators of the square-law type must be small, so that the current never goes quite to zero at any point in the cycle, if true square-law operation is to be realized. As long as this is true, the performance may be readily analyzed by using the power-series analysis of Chap. 12, using only the first two terms of Eq. (12-5)[1] As will be shown, the modulated wave

[1] The power-series analysis is applicable to all types of operation but, if the plate current is zero for an appreciable portion of the cycle, satisfactory results are obtainable only by using so many terms of the series as to make this method of analysis very laborious.

is produced by the *second-order* term in this equation, and it is for this reason that this type of modulation is known as *square-law*.

Analysis of the Performance of a Square-law Modulator. The performance of the square-law modulator may be determined by writing the equation of the current from Eq. (12-5) and then multiplying each component of this current by the load impedance presented at the frequency of the component. This will give the equation of the output voltage which will, perhaps, be applied to the grid of the next amplifier tube. Let us assume a r-f amplifier such as that of Fig. 10-3, page 387, operating on a curved portion of its characteristic curve. Let the carrier signal have a frequency of $\omega/2\pi$ and let some modulating signal be introduced having a frequency of $q/2\pi$, where q is very much less than ω. This means that the signal voltage applied to the nonlinear amplifier is that of Eq. (12-49), and therefore, the plate current is equal to $I_b - i_{p0}$ where i_{p0} is given by Eq. (12-52). The voltage developed across the output impedance of the amplifier is then found by multiplying each term of the plate current by the impedance presented at its frequency.

Let the impedance presented by the load circuit be R_L at resonance ($f = \omega/2\pi$) and zero at all frequencies appreciably different from that of resonance, a condition closely realized in practice. Then those components of current having angular frequencies of ω, $\omega + q$, and $\omega - q$ will encounter a load resistance of R_L and will produce an output voltage, while the remaining components will encounter zero-load impedance and will therefore produce no voltage. The angular frequencies $\omega + q$ and $\omega - q$ are normally so close to the resonant frequency ω that the load impedance at these frequencies may be taken as R_L.

Under the foregoing assumptions the output voltage is equal to

$$e = a_{1(\omega)}R_LE_{1m}\sin\omega t - a_{2(\omega+q)}R_LE_{1m}E_{2m}\cos(\omega+q)t + a_{2(\omega-q)}R_LE_{1m}E_{2m}\cos(\omega-q)t \quad (13\text{-}18)$$

When a_1 and a_2 are evaluated from Eqs. (12-46) and (12-47), assuming constant μ, as in a diode or in many triodes, the result is

$$e = -\frac{\mu R_L E_{1m}}{r_p + R_L}\sin\omega t - \frac{\mu^2 r_p' R_L E_{1m}E_{2m}}{2(r_p+R_L)^2}\cos(\omega+q)t + \frac{\mu^2 r_p' R_L E_{1m}E_{2m}}{2(r_p+R_L)^2}\cos(\omega-q)t \quad (13\text{-}19)^*$$

The first term in this equation is the carrier, and the other two are the side bands

Equation (13-5) shows the magnitude of each side band to be $m/2$ times that of the carrier if the modulation is distortionless; i e., $m/2 = E_{sb}/E_{\text{carrier}}$. Therefore, we may write from Eq (13-19)

$$\frac{m}{2} = \frac{\mu^2 r_p' R_L E_{1m} E_{2m}/2(r_p + R_L)^2}{\mu R_L E_{1m}/(r_p + R_L)}$$

$$m = \frac{\mu r_p' E_{2m}}{r_p + R_L} \qquad (13\text{-}20)^*$$

Two conclusions may be drawn from Eq (13-20). (1) The percentage of modulation achieved is proportional to the amplitude of the modulating voltage E_{2m} regardless of the intensity of the carrier A study of Fig 13-18 will show that this might have been expected, since the amplitude of the alternating current flowing depends on the slope of the $i_b - e_b$ curve which in turn is a function of the instantaneous value of the modulating voltage (the curve marked $E_{2m} \sin qt$ in Fig. 13-18). (2) The percentage of modulation is comparatively low. The derivative of the plate resistance (r_p') is much smaller than r_p for most tubes, so that even a large modulating voltage produces only a low level of modulation Since most square-law modulation occurs under conditions where it is undesired, as in the hum-modulation example suggested in a previous paragraph, even a low level of modulation is objectionable

If pentode tubes are used and the load impedance is sufficiently low compared to r_p so that the approximate relations of Eqs. (12-10) and (12-24) may be used, Eq (13-18) becomes[1]

$$e = -g_m R_L E_{1m} \sin \omega t + \frac{g_{mc}' R_L E_{1m} E_{2m}}{2} \cos (\omega + q)t - \frac{g_{mc}' R_L E_{1m} E_{2m}}{2} \cos (\omega - q)t \quad (13\text{-}21)^*$$

and the modulation factor becomes

[1] It may appear, from a comparison of this equation with Eq (13-19), that the phase of the side frequencies with pentode tubes is opposite to that obtained with triodes But r_p' is a negative number since r_p *decreases* with increasing plate voltage whereas g_{mc}' is a positive number since g_m *increases* with increasing grid voltage. Thus both equations give the same relative phase for each side frequency.

$$m = \frac{g'_{mc} E_{2m}}{g_m} \tag{13-22}^*$$

Again, m is small compared to unity, since g'_{mc} is normally small compared to g_m.

Components of Square-law Modulation. A square-law modulator evidently functions essentially as an amplifier that is operated on the curved portion of its characteristic curve and so produces amplitude distortion. Thus the discussion in the section beginning on page 505 is applicable to square-law modulators. By applying this approach it is possible to predict the frequency of all components to be found in the output of the modulator without first obtaining Eq. (13-19). In the example of the preceding section the plate current must contain the following components: ω, q, $\omega + \omega = 2\omega$, $\omega - \omega = 0$, $q + q = 2q$, $q - q = 0$, $\omega + q$, $\omega - q$. The output circuit filters out all these components except ω, $\omega + q$, and $\omega - q$. If the characteristic curve of the modulator tube is not truly square-law in shape, third-order terms may also be present including $2\omega + \omega = 3\omega$, $2\omega - \omega = \omega$, $2q + q = 3q$, $2q - q = q$, $2\omega + q$, $2\omega - q$, $2q + \omega$, and $2q - \omega$. These are also filtered out by the output circuit.

Cross Modulation. As stated in a preceding section, *any* nonlinear circuit element is capable of producing square-law modulation. Thus square-law modulation of radio signals may occur under conditions and in circuits where it is definitely not desired. Such undesired production, whether in a vacuum tube or other nonlinear impedance, is commonly known as *cross modulation*. More specifically cross modulation is defined in the 1938 Standards Report of IRE as "a type of intermodulation due to modulation of the carrier of the desired signal by an undesired signal."

A common source of cross modulation is found in the first amplifier of a radio receiver. Signals from several radio transmitting stations are normally impressed on the receiving antenna simultaneously and therefore on the input circuit to the first tube. The selective circuits used in this input circuit are incapable of completely eliminating all undesired signals (the additional selectivity of successive amplifier stages being required to complete the separation); and unless the first tube is operated on the straight portion of its characteristic curve, square-law demodulation and modulation will take place. If the incoming signals are from

radiotelephone stations (*e.g.*, broadcasting stations), the modulation carried by each station will be reproduced as an audio signal by demodulation (see next chapter) and will be remodulated on any carrier waves present in the circuit. Thus if the receiver is tuned to station A and the signal from station B is sufficiently strong to produce a small but appreciable voltage on the grid of the first tube, the modulation of station B will also be impressed on the carrier of station A. No degree of selectivity in the remaining circuits of the receiver will then have any effect upon this superimposed modulation, and it will be heard in the output of the receiver as an interference signal. Thus it is very important that the first tube (and to a lesser extent successive tubes) in a radio receiver shall not operate on the curved portion of its characteristic curve. This requirement made necessary the development of the remote cutoff tubes (page 82) before automatic volume control[1] could be successfully employed, the characteristic curves of these tubes being essentially straight, for small signal voltages, no matter what bias is applied.

Cross modulation may also occur in nonlinear impedances external to the receiver, causing a carrier wave that already carries the modulation of two or more different stations to be impressed on the receiver input. The only possibility of eliminating this form of cross modulation is to locate the antenna at a point remote from any such source. Possible sources are power and telephone wires, water pipes, or other conducting material. The nonlinear impedances found in these conducting materials are produced by discontinuities such as poor splices in electric wiring or water pipes touching each other. Experience seems to indicate that in a power distribution system even a splice capable of safely passing large amounts of 60-cycle power may have sufficient nonlinearity to cause cross modulation.

Another effect of such cross modulation is to produce new carrier waves by interaction of two or more powerful radio transmitting stations. In Eq. (12-52) the expansion of Eq. (12-5) was carried out to include first- and second-order terms only. If extended to include third-order terms[2] the components $(2\omega - q)$ and $(2q - \omega)$ (where ω and q represent the carrier frequencies of two different transmitters) very commonly lie in the tuning range

[1] See p. 587 for a discussion of automatic volume control action.

[2] See discussion on p. 506.

of the radio receiver. Thus if the receiver is tuned to this cross-modulation (or spurious) frequency, a signal will be received carrying the modulation of both transmitters. The extent of this phenomenon is, of course, a function of the signal strength of the two stations as well as of the presence of non-linear impedances in proximity to the receiving antenna.[1]

Balanced Modulator. It is sometimes desirable to remove the carrier from a modulated wave, leaving only the two side bands. This may be done by means of the balanced modulator of Fig. 13-20. This modulator consists of two r-f amplifiers operated push-pull for the modulating signal and in parallel for the carrier, the tubes commonly operating class A as van der Bijl, square-law modulators.

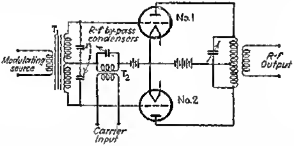


FIG. 13-20. Circuit of a balanced modulator. If the carrier and modulating-input circuits to the tube are interchanged, it becomes a push-pull, grid modulated amplifier.

Let the voltage in the upper half of the secondary of transformer T_1 be E_{am} sin qt; then that in the lower half will be $-E_{am}$ sin qt. Let the carrier voltage be E_{0m} sin ωt. Then the alternating voltage applied to the grid of tube 1 will be

$$e_{g_1} = E_{0m} \sin \omega t + E_{am} \sin qt \tag{13-23}$$

and that applied to tube 2,

$$e_{g_2} = E_{0m} \sin \omega t - E_{am} \sin qt \tag{13-24}$$

The plate current flowing in each tube may be determined by in-

[1] For more details concerning this problem, see Austin V. Eastman and Lawrence C. F. Horle, The Generation of Spurious Signals by Non Linearity of the Transmission Path, *Proc. IRE*, **28**, p. 438, October, 1940.

serting Eqs. (13-23) and (13-24) in Eq. (12-5), page 190. The incremental plate current for tube 1 is then

$$i_{p0_1} = a_1 E_{0m} \sin \omega t + a_1 E_{am} \sin qt + \frac{a_2 E_{0m}^2}{2} - \frac{a_2 E_{0m}^2}{2} \cos 2\omega t + a_2 E_{am} E_{0m} \cos (\omega - q)t - a_2 E_{am} E_{0m} \cos (\omega + q)t + \frac{a_2 E_{am}^2}{2} - \frac{a_2 E_{am}^2}{2} \cos 2qt + \cdots \qquad (13\text{-}25)$$

The incremental plate current for tube 2 may be determined in the same manner and will be

$$i_{p0_2} = a_1 E_{0m} \sin \omega t - a_1 E_{am} \sin qt + \frac{a_2 E_{0m}^2}{2} - \frac{a_2 E_{0m}^2}{2} \cos 2\omega t - a_2 E_{am} E_{0m} \cos (\omega - q)t + a_2 E_{am} E_{0m} \cos (\omega + q)t + \frac{a_2 E_{am}^2}{2} - \frac{a_2 E_{am}^2}{2} \cos 2qt + \cdots \qquad (13\text{-}26)$$

The net magnetization in the output transformer will be set up by the difference in the NI produced by the two plate currents. Evidently, in taking the difference, all components of the two currents will cancel out[1] except $\omega + q$, $\omega - q$, and q. The q term will not be repeated through the output transformer, however, since the tank circuit is tuned to ω, where $\omega \gg q$. The output will, therefore, contain only the two side bands.

If the carrier and modulating sources are interchanged in the circuit of Fig. 13-20, the resulting plate current in the two tubes will be given by Eqs. (13-25) and (13-26) by interchanging q and ω. The additive terms will then be ω, $\omega + q$, and $\omega - q$,* giving the usual modulated wave. Such a modulator is merely a push-pull, square-law, modulated amplifier or, if the tubes are biased for class C operation, a push-pull, linear, grid-modulated amplifier.

Single Side-band Transmission. Elimination of the carrier, as in a balanced modulator, is desirable, as 67 per cent of the power transmitted, with 100 per cent modulation, is contained in the carrier. Since the carrier contains no information, it is much more

[1] Assuming a_1 and a_2 to be the same for both tubes, i.e., the tubes must have identical characteristics.

* This term should be $q - \omega$, but $\cos (q - \omega) = \cos (\omega - q)$ so that it seems more logical to write the higher frequency term first.

economical to supply it locally at the receiver than to supply it from the transmitter. Unfortunately the carrier must be reinserted at the receiver with an accuracy that is practically impossible to realize. As will be shown in Chap. 14, each side band, when combined with the carrier, produces a component of current of the modulating frequency q, as desired. If the carrier is reinserted with exactly the same phase relative to the side bands as it had in the original transmission, these two components will be additive; if the phase of the carrier is shifted 90 deg, they will cancel. Any other phase shift will produce only partial addition of the two components. Thus a difference in frequency between the original carrier and the reinserted carrier of only a fraction of a cycle per second will produce a continuously varying phase difference between the carrier and side frequencies, and the output will vary periodically between maximum and zero. Even if the reinserted carrier is of exactly the same frequency as the original there is no assurance that its phase will be such as to produce maximum output.[1]

If one of the side bands, as well as the carrier, is removed, the problem of reinserting the carrier in the correct phase is eliminated, since only one of the two modulating-frequency components is then present in the detector output; furthermore only half the band width is then needed to transmit a given intelligence. Removal of one side band may be achieved by the use of filters that pass only the frequencies in one of the two side bands. Since the degree of separation realizable with practical filters is a function of the ratio of the desired to the undesired frequencies, it is far easier to perform this separation at a low than at a high frequency. Suppose, for example, that a carrier of 10 kc is modulated by a band of frequencies lying between 100 and 3000 cycles. If a balanced modulator is used, the output contains components in the frequency bands 7000 to 9900 and 10,100 to 13,000, the carrier being removed. It is quite feasible to separate these two bands by conventional band-pass filters and so remove one, say the lower. The remaining upper side band may now be used to modulate a carrier of, say, 90 kc and the two new side bands are 77,000 to 79,900 and

[1] In one solution of this problem a small part of the carrier is transmitted and, at the receiver, either is used to synchronize a local oscillator or, by means of crystal filters, is separated from the side bands and then amplified and reinserted. Such methods will hold the carrier in the correct phase.

100,100 to 103,000. Again these two groups may be separated, and one of them used to modulate a higher carrier frequency. This procedure may be carried as far as desired to modulate a carrier wave of any frequency and so transmit a single side band and no carrier.[1]

2. FREQUENCY AND PHASE MODULATION

When either frequency or phase modulation is used, the amplitude of the transmitted wave remains constant, but either its frequency or its phase is varied at a rate proportional to the audio modulating frequency and by an amount proportional to the amplitude of the audio signal. The two methods are really very similar, since no change in frequency can occur without at least a momentary shift in phase, nor can a change in phase occur without a momentary change in frequency. Consequently, the resulting waves and equations for these two types are almost identical.

Let us assume that a carrier wave of 100 Mc is to be frequency-modulated by an audio signal of 800 cycles/sec. With a certain intensity of modulating signal the transmitted frequency might vary 10,000 cycles/sec either side of the unmodulated carrier frequency, or from 99.99 to 100.01 Mc, the variation being at an 800 cycles/sec rate. If the intensity of the modulating signal is increased, the frequency would deviate by more than 10,000 cycles/sec from the original carrier frequency, and if the frequency of the modulating signal is increased, the rate of variation of the radio signal will increase accordingly. Thus, in a f-m wave,[2] the *amplitude* of the modulating signal is indicated by the *amount* of frequency deviation and the *frequency* of the modulating signal is indicated by the *rate* of frequency deviation of the radio wave that is being modulated.

All the statements of the preceding paragraph may be applied equally well to a phase-modulated wave by simply substituting the word "phase" for "frequency" in the appropriate places.

Analysis of Frequency and Phase Modulation. The equations

[1] See also N. Koomans, Single-side-band Telephony Applied to the Radio Link between the Netherlands and the Netherlands East Indies, *Proc. IRE*, 26, p. 182, February, 1938; and C. T. F. van der Wyck, Modern Single-side-band Equipment of the Netherlands Postals Telephone and Telegraph, *Proc. IRE*, 36, pp. 970–980, August, 1948.

[2] "f-m" is a commonly used abbreviation for "frequency-modulated."

for frequency- and phase-modulated waves may be determined as follows:[1] An alternating voltage may be represented by

$$e = E_m \sin \varphi \tag{13-27}$$

where φ is the instantaneous angular displacement of the voltage vector from the zero axis. Since ω is the angular velocity of the voltage vector, we may also write

$$\omega = \frac{d\varphi}{dt} \tag{13-28}$$

from which it is possible to solve for φ by integration

$$\varphi = \phi + \int \omega \, dt \tag{13-29}$$

where ϕ is the integration constant and represents the initial phase displacement.

In both amplitude and phase modulation ω may be considered constant and equal to ω_0; therefore, the integration indicated in Eq. (13-29) reduces to simply $\omega_0 t$. Equation (13-27) may then be written in the more familiar form

$$e = E_m \sin (\omega_0 t + \phi) \tag{13-30}$$

When amplitude modulation is applied, it is necessary only to let E_m vary with the modulating signal which, if ϕ is assumed to be zero, results in Eq. (13-3). For phase modulation, however, the amplitude remains constant, but the phase angle may be written as

$$\phi = \phi_0 + m_p \sin qt \tag{13-31}$$

where m_p is the maximum phase deviation from the initial phase ϕ_0 and is a function of the amplitude of the modulating signal. When this is substituted in Eq. (13-30) and the initial phase angle ϕ_0 is assumed to be zero, the result is

$$e = E_m \sin (\omega_0 t + m_p \sin qt) \tag{13-32*}$$

which is the equation for a phase-modulated wave, comparable to Eq. (13-3) for a-m waves.

[1] Hans Roder, Amplitude, Phase and Frequency Modulation, *Proc. IRE*, 19, p. 2145, December, 1931; also see Balth van der Pol, Frequency Modulation, *Proc. IRE*, 18, p. 1194, July, 1930.

The equation for f-m waves may be found by writing[1]

$$f = f_0 + kf_0 \cos qt \tag{13-33}$$

or

$$\omega = \omega_0 + 2\pi k f_0 \cos qt \tag{13-34}$$

where kf_0 is the maximum frequency deviation, k being a function of the amplitude of the modulating signal, and f_0 is the carrier frequency. This equation must be inserted into Eq. (13-29) and the indicated integration carried out. If the value of φ so obtained is inserted into Eq. (13-27) and the initial phase angle ϕ is taken as zero, the result is

$$e = E_m \sin (\omega_0 t + m_f \sin qt) \tag{13-35}^*$$

where $m_f = kf_0/f_q$
$f_q = q/2\pi$

It is evident from Eqs. (13-32) and (13-35) that frequency and phase modulation are similar, both producing an instantaneous phase shift that is proportional to the instantaneous amplitude of the modulating signal.[2] However, the instantaneous phase shift resulting from frequency modulation is also inversely proportional to the *frequency* of the modulating signal. Herein lies the essential difference between phase and frequency modulation.[3]

It should be evident from the foregoing that it is possible to produce phase modulation by frequency-modulating the wave according to Eq. (13-33) provided only that the frequency deviation kf_0 varies directly with the frequency of the modulating signal as well as with its amplitude. Conversely, it is possible to produce a f-m wave by varying the phase according to Eq. (13-31) provided only that m_p varies inversely with the frequency of the modulating signal as well as directly with its amplitude. Either of these conditions may be realized in practice by the use of suitable filters inserted in series with the modulating signal.

Equations (13-32) and (13-35) may be expanded only with the

[1] cos qt is used here, rather than sin qt as in Eq. (13-31), to give a final equation of the same form as Eq. (13-32).

[2] Comparison of Eqs. (13-32) and (13-35) with Eq. (13-30) shows that $m_p \sin qt$ in the first equation and $m_f \sin qt$ in the second are comparable in effect to ϕ and may therefore be considered as producing a phase shift that varies at a frequency $q/2\pi$.

[3] See also Herbert J. Scott, Frequency vs. Phase Modulation, *Communications*, 20, p. 10, August, 1940.

aid of Bessel's functions. As many engineers are unfamiliar with these functions, the results only of such an expansion are given in the following equation:[1]

$$\begin{aligned} e = E_m\{J_0(x) \sin \omega_0 t &+ J_1(x)[\sin (\omega_0 + q)t - \sin (\omega_0 - q)t] \\ &+ J_2(x)[\sin (\omega_0 + 2q)t + \sin (\omega_0 - 2q)t] \\ &+ J_3(x)[\sin (\omega_0 + 3q)t - \sin (\omega_0 - 3q)t] \\ &+ \cdots\} \end{aligned} \quad (13\text{-}36)^*$$

in which $J_0(x)$, $J_1(x)$ . . . are Bessel's functions of x. The factor x is equal to m_p for phase modulation and to m_f for frequency modulation, m_p and m_f being known as the modulation indexes.

The first term of Eq. (13-36) is the carrier, and the remaining terms are side frequencies, of which there are an infinite number. The amplitude of all components varies with the modulation index, being equal to the product of the amplitude of the unmodulated carrier and the appropriate Bessel's function as determined from Fig. 13-21.[2] Evidently, if the modulation index is less than about 0.5, only the first side band is of much importance, whereas additional side bands become of importance as the index is increased. Evidently, too, since m_f varies inversely as the audio frequency, frequency modulation will produce a large number of side bands

[1] The following relations hold true:

$$\sin (x \sin r) = 2J_1(x) \sin r + 2J_3(x) \sin 3r + \cdots$$

$$\cos (x \sin r) = J_0(x) + 2J_2(x) \cos 2r + 2J_4(x) \cos 4r + \cdots$$

where $J_0(x)$ is given by

$$J_0(x) = 1 - \frac{x^2}{2^2} + \frac{x^4}{2^2 \cdot 4^2} - \frac{x^6}{2^2 \cdot 4^2 \cdot 6^2} + \cdots$$

and $J_n(x)$ is given by

$$J_n(x) = \frac{x^n}{2^n n!}\left[1 - \frac{x^2}{2(2n+2)} + \frac{x^4}{2 \cdot 4(2n+2)(2n+4)} - \frac{x^6}{2 \cdot 4 \cdot 6(2n+2)(2n+4)(2n+6)} + \cdots\right]$$

where n is an integer. When these expansions are inserted in Eq. (13-32) or (13-35), the result may be simplified to give Eq. (13-36). Computed values of $J_0(x)$ and $J_n(x)$ are given in Fig. 13-21.

[2] See Roder, *loc. cit.*; also Earl D. Scott and John R. Woodyard, Side Bands in Frequency Modulation, *Univ. Wash. Eng. Exp. Sta. Bull.* 68, 1932.

when the modulation frequency is low but comparatively few when it is high This is not true of phase modulation, since m_p is independent of f_q.

Band Width. An important factor in the design of a radio transmitter or receiver is the range of frequencies that must be transmitted to reproduce the modulated wave. For proper reproduction all the side bands as well as the carrier must be amplified equally, and for amplitude modulation the circuit must therefore provide reasonably uniform response over a range of

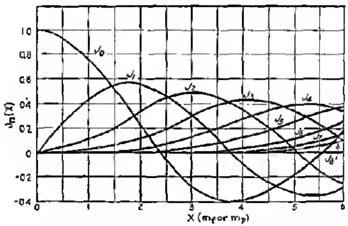


FIG. 13-21. Curves of the first nine Bessel functions

frequencies from $f_0 - f_q$ to $f_0 + f_q$ or the band width is equal to $(f_0 + f_q) - (f_0 - f_q) = 2f_q$, where f_q is the highest frequency component in the modulating signal and f_0 is the carrier frequency. In frequency modulation, on the other hand, the band of frequencies to be transmitted extends from $(f_0 - nf_q)$ to $(f_0 + nf_q)$, where n is equal to the number of side bands of sufficient magnitude to be of importance, and the band width is therefore $2nf_q$. Roughly speaking, the highest order side band which is of importance is about $(m_f + 1)$, all higher order terms being less than 15 per cent of the unmodulated carrier, as may be checked from Fig. 13-21 for small values of m_f. If we designate band width as bw, we may now write the two following equations:

$$bw = 2nf_q \tag{13-37}$$

$$n = m_f + 1 \tag{13-38}$$

Substituting Eq. (13-38) into (13-37) gives

$$bw = 2(m_f f_q + f_q) \tag{13-39}$$

From the expression for m_f immediately following Eq. (13-35) we may write

$$m_f f_q = kf_0 \tag{13-40}$$

Substituting Eq. (13-40) into (13-39) gives for the band width

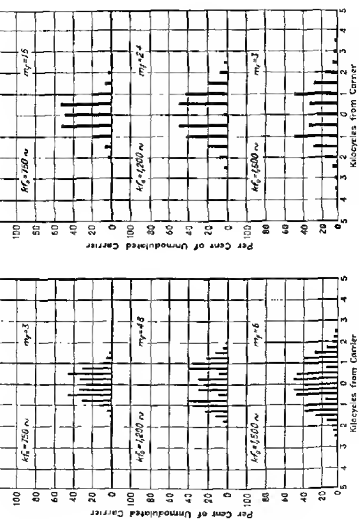
$$bw = 2(kf_0 + f_q) \tag{13-41*}$$

This equation indicates that, if kf_0 is much larger than f_q as is usually the case, the band width in a f-m wave is essentially proportional to the amplitude of the modulating signal (since k is proportional to this amplitude) and is very much larger than that of an a-m wave.

The significance of Eq. (13-41) may be more clearly seen with the aid of Fig. 13-22,[1] where the side frequencies and carrier, calculated from Eq. (13-36) for various values of kf_0 and f_q, are indicated by heavy vertical lines. The vertical scale indicates the magnitude, and the horizontal scale, the frequency relative to the unmodulated carrier. Only small values of kf_0 have been chosen since the number of side frequencies present at the usual large values of kf_0 makes picturization impossible. A check of Eq. (13-41) against Fig. 13-22 will show that all side frequencies more than $(kf_0 + f_q)$ either side of the carrier have an amplitude of less than 15 per cent of the carrier, thus corroborating the conclusions drawn from the equation. Note particularly that whenever m_f is less than about 0.5 the carrier amplitude is virtually equal to its unmodulated value, a fact which is of importance in the analysis of the Armstrong system of modulation presented in a later section.

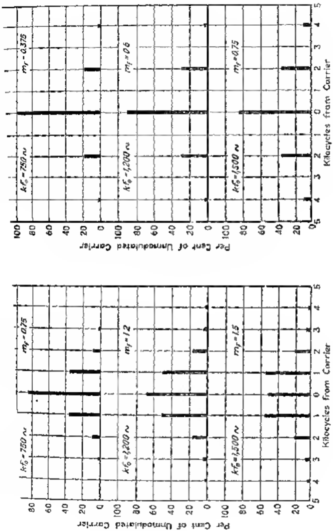
While the very wide band width required for large values of kf_0 presents problems in designing radio equipment, it results in a number of improvements in transmission as demonstrated in the next section. The problems of equipment design are usually solved by increasing the carrier frequency until the ratio of band

[1] Adapted from Scott and Woodyard, *op. cit.*



(a) $f_a = 250$ cps

(b) $f_a = 500$ cps



(c) $f_q = 1{,}000$ c.p.s.

(d) $f_q = 2{,}000$ c.p.s.

Fig. 13-22. Side-band frequency spectra in a f-m wave.

width to carrier frequency is sufficiently low so that resonant circuits may be designed to pass all the side bands and yet be satisfactorily selective. Carrier frequencies for use with frequency modulation range from about 30 Mc on up

For phase modulation the problem is somewhat different since the modulation index m_p is independent of the frequency of the modulating signal. Replacing m_f with m_p in Eq. (13-39) gives, for phase modulation,

$$bw = 2(m_p f_s + f_s) \qquad (13\text{-}42)^*$$

This indicates that the band width, if m_p is considerably greater than unity (the usual case), is essentially proportional to the frequency of the modulating signal as well as to its amplitude (m_p being proportional to the amplitude) so that the maximum band width to be designed for is that required to transmit the highest audio frequency at the maximum amplitude, at other modulating frequencies only a portion of the available band width will be used. This does not permit full realization of the benefits of noise reduction as given in the next section, and phase modulation is used less than frequency modulation.

Analysis by Vector Diagrams. A principal advantage of frequency and, to a lesser extent, of phase modulation over amplitude modulation is that the signal-to-noise ratio is much higher with the former than with the latter. To understand the process of noise reduction it is desirable to study all three types of modulation from the standpoint of vector diagrams, although, since phase and frequency modulation differ only in the *degree* of modulation, owing to the factor f_s, only amplitude and frequency modulation will be considered in the ensuing discussion.

An unmodulated carrier wave may be represented by a vector whose length is equal to the crest value of the wave, which is rotating at an angular velocity ω. If the reader can imagine himself rotating at the same angular velocity, this vector would appear to stand still as in Fig. 13-23a.[1] If the wave is now amplitude-modulated, the length of the vector will vary but its position will remain unchanged. With 100 per cent modulation the vector will pulsate between a maximum of twice the original

[1] For those whose imagination will not permit this trick a similar result may be achieved by considering the carrier vector to be a physical entity, illuminated by stroboscopic light of the same frequency as the wave.

carrier, as shown by the dotted line in Fig. 13-23*b*, and a minimum of zero. With frequency modulation, on the other hand, the magnitude will remain constant while the frequency and, therefore, the phase vary periodically as in Fig. 13-23*c*. In actual practice this shift in phase may be as much as many times 2π radians; in other words the vector may rotate many times in a clockwise direction and then return and rotate an equal number of times in a counterclockwise direction.

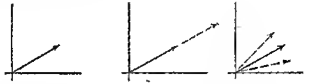


FIG. 13-23. Vectors illustrating (*a*) an unmodulated carrier wave, (*b*) an a-m wave, (*c*) a f-m wave.

The same conclusions may be reached by bringing in the concept of side frequencies. An a-m wave may then be represented by the sum of three vectors, each rotating at different angular velocities, *viz.*, at ω, at $\omega + q$, and at $\omega - q$. These are represented in Fig. 13-24*a* where the side-frequency vectors are half as long as the carrier vector, as in 100 per cent modulation. If the reader again imagines himself rotating at the same velocity as the carrier vector, this vector will seem to stand still, while one side-frequency vector will rotate clockwise at an angular velocity q and the other will rotate in a counterclockwise direction at the same velocity.

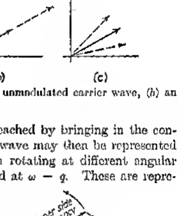


FIG. 13-24. Vectors illustrating the relative phase of carrier and side bands in an a-m wave.

For most purposes it is better to place the side-frequency vectors at the end of the carrier vector as in Fig. 13-24*b*, whereupon the resulting vector is more readily found by taking the sum of the three. It should be quite evident that as the two side-frequency

vectors rotate in opposite directions at an angular velocity q, their sum is another vector which is always in phase with the carrier but which varies sinusoidally in magnitude between positive and negative limits that are equal to the carrier for 100 per cent modulation, giving a resultant total that varies between zero and twice the carrier. When plotted in rectangular coordinates this, of course, gives a wave similar to that of Fig. 13-2.

A f-m wave may be represented in the same manner, but in the interest of simplicity it will be assumed that the degree of modulation is so low that only the first-order side frequencies are of importance. Under this assumption a f-m wave may be represented by the vectors of Fig. 13-25. This diagram is distinguished from that of the a-m wave in that the two side frequencies are displaced by 90 deg, relative to the carrier, from their positions in Fig. 13-21. (This 90-deg relationship is shown in Eq. (13-36) where the first-order side frequencies are proportional to the *sine* of the angle whereas in Eq. (13-5) for amplitude modulation the side-frequency terms are proportional to the *cosine* of the angle.) Evidently the sum of the two side frequencies in Fig. 13-25 is a vector that lies at right angles to the carrier vector and, when added to the carrier, produces a resultant that shifts in phase but is essentially constant in magnitude.[1]

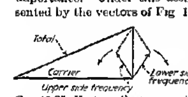


FIG. 13-25. Vectors illustrating the relative phase of carrier and first-order side bands in a f-m wave.

Noise and Interference Reduction. The manner in which frequency modulation reduces noise and other interference may now be seen with the aid of Fig. 13-26. If the carrier wave is for the moment assumed to be unmodulated, the sum of this carrier and of the noise or other interfering signal will be the voltage impressed on the receiver, the locus of which is the circle in Fig. 13-26. (The noise vector will in general vary in both phase and magnitude, although phase shift only is indicated in the figure.)

[1] The resultant should of course be exactly constant in magnitude as in Fig. 13-23c. The reason it is not is that only the first order side frequencies were taken into consideration and the carrier was considered as constant in amplitude. If the effect of higher order side bands is considered and the carrier vector varies according to $J_0(m_f)$, the resultant will be found constant in length throughout the modulation cycle.

Since the receiver will respond to variations in phase (or frequency) only, the noise output from the receiver will evidently be proportional to the angular shift ϕ which, if the magnitude of the interference is not more than about half that of the desired signal, will not exceed $\frac{1}{2}$ radian.

If the desired signal is now frequency-modulated so as to produce a maximum phase shift of many radians, the ratio of signal to noise in the output of the receiver will be very high. This follows since the signal-to-noise ratio is proportional to the ratio of the phase shift due to the desired modulation to that due to the noise, and the latter is limited to $\frac{1}{2}$ radian for a noise signal 50 per cent of the magnitude of the desired carrier. There is no limit to the amount of desired modulation that may be applied, as there is in amplitude modulation where m must not exceed 1; although an increase in band width always accompanies an increase in the frequency deviation, as indicated by Figs. 13-21 and 13-22, where additional side frequencies are seen to become of appreciable magnitude whenever m_f is increased.

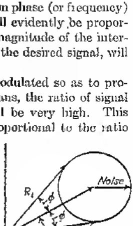


FIG. 13-26. Vector diagram illustrating the manner in which frequency modulation reduces noise.

From the foregoing analysis based on Fig. 13-26 a large frequency deviation and, therefore, a wide band of frequencies are apparently required for effective noise reduction. Within limits this is true, and modern f-m transmitters are operated with a band width of about 200 kc. However, any increase in band width is accompanied by an increase in the noise and other interfering signals that fall within the response band of the receiver. Thus the improvement in noise reduction, as the band width is increased, is not quite so great as might otherwise be expected.

Since the noise vector shown in Fig. 13-26 might just as well represent an interfering station, it is evident that no serious interference may be expected from another transmitting station on the same frequency as the desired wave unless the amplitude of the interfering signal is greater than about 50 per cent of the desired one. Thus it is possible to operate f-m transmitting stations on the same frequency at locations much closer together than with a-m stations.

Methods of Producing Frequency Modulation. The simplest method of producing frequency modulation is to vary the capacitance of the tank circuit of the oscillator by means of a small condenser, the capacitance of which is capable of being varied by the modulating source. It is, however, difficult to secure a good response from such a device over the entire audio spectrum

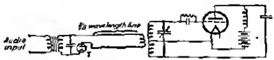


FIG. 13-27. Circuit of a diode modulator together with an eighth-wave line for producing frequency modulation

Another method is to use an eighth-wave transmission line as shown in Fig. 13-27.[1] With no modulation voltage applied, the transmission line is short-circuited half the time and open-circuited half the time, provided the impedance of the diode is negligible on the positive half cycle. This is indicated in Fig 13-28 between points a and b where the alternating voltage is the carrier and the (+) and (−) signs indicate the polarity of the voltage applied to the diode Since a low-loss eighth-wave transmission line presents a nearly pure inductive reactance when short-circuited and a nearly pure capacitive reactance of equal magnitude when open-circuited, the average reactance under no-modulation conditions is zero When modulation voltage is applied, the line is short-circuited more of the time than it is open-circuited during one half cycle of the modulation voltage, as from b to c, and presents a net inductive reactance, thus shifting the frequency. On the next half cycle, the line presents a net capacitive reactance, as

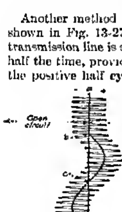


FIG 13 28 Illustrating the manner in which the circuit of Fig 13 27 produces frequency modulation.

[1] Austin V. Eastman and Earl D. Scott, Transmission Lines as Frequency Modulators, *Proc. IRE*, **22**, p 878, July, 1934. The circuit of Fig. 13 27 represents an improvement over that given in the reference, suggested by John Woodyard

from c to d, and the frequency is shifted in the opposite direction, producing a f-m wave.

Balanced Modulator System of Frequency Modulation. Neither of the two methods outlined in the preceding section includes any provision for maintaining the average (or carrier) frequency exactly at the assigned value. As it is impossible to frequency-modulate a crystal oscillator, the problem of frequency stability is one that offers serious complications in any system involving frequency modulation.

Major Armstrong solved this problem by starting out with amplitude modulation of the output of a crystal oscillator and then converting this signal into a phase-modulated wave.[1] The

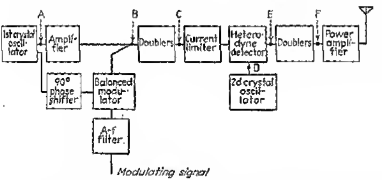


FIG. 13-29. General arrangement of Major Armstrong's f-m transmitter.

general arrangement of his circuit is indicated in the block diagram of Fig. 13-29. The output of a crystal oscillator is fed to a balanced modulator (see page 533) through a network that shifts the phase by 90 deg. The output of the balanced modulator contains only the side bands, and these are then combined with the carrier in its original phase. The result is a carrier and pair of side bands related as in Fig. 13-25 which, if the phase shift is kept small (not to exceed about 0.5 radian), is a very close approximation to a phase-modulated wave.[2] Figure 13-22 shows

[1] Edwin H. Armstrong, A Method of Reducing Disturbances in Radio Signaling by a System of Frequency Modulation, *Proc. IRE*, 24, p. 689, May, 1936.

[2] Note that the approximation is to a *phase*-modulated wave, not to a *frequency*-modulated wave, since the modulation amplitude is independent of the frequency of the modulating wave.

that for $m = 0.5$ or smaller only the first-order side bands in a phase- or frequency-modulated wave are significant and the carrier is nearly equal to its no-modulation value. Thus the conditions of Fig. 13-22*d* for $m = 0.375$ are almost exactly realized by Armstrong's system which produces a 100 per cent carrier and first-order side bands only.

Three objections to this method are: (1) the amplitude is not absolutely constant as it should be,[1] (2) phase modulation and not frequency modulation is produced, (3) the phase shift is not sufficient to produce the desired improvement in signal-to-noise ratio. The first objection is overcome easily by inserting a current-limiting unit, such as an overdriven class C amplifier, in series with the output (Fig. 13-29). Since the crest value of the alternating plate voltage in a class C amplifier cannot exceed the direct voltage, it is evident that the output of the current-limiting amplifier cannot increase beyond a certain point no matter how great may be the grid excitation. Amplitude distortion that may be generated in this stage is not objectionable, since the receiver is to respond to changes in frequency rather than to changes in amplitude. Insertion of the limiter will automatically supply the necessary higher order side bands and adjust the carrier amplitude to conform to Eq. (13-36) and Fig. 13-22.

The second objection is raised because phase modulation will not give so much noise reduction as will frequency modulation. It was shown in a preceding section that effective noise reduction requires large values of m, and phase modulation is incapable of producing so large an *average* value of m as is frequency modulation for the same maximum band width. This may be seen by noting that Eq. (13-42) is nearly equal to

$$bw = 2m_pf_s \tag{13-43}$$

if m_p is much larger than one. Also m_p is proportional to the amplitude of the modulating signal or

$$m_p = k_pE_a \tag{13-44}$$

[1] This is obvious from a study of Fig. 13-25. For the amplitude to be constant, higher order side bands should be present, and the carrier amplitude should decrease with modulation, as shown by the $J_0(x)$ curve of Fig. 13-21.

where k_p is a proportionality factor and E_a is the amplitude of the modulating signal. Inserting Eq. (13-44) into (13-43) gives

$$bw = 2k_pE_af_q \tag{13-45}$$

which shows that for a given amplitude the band width is proportional to the frequency of the modulating signal. Thus at low modulating frequencies the band width will be small, the maximum band width being required at the highest value of f_q. If we assume that a radio transmitter is designed to amplify a band width of 200,000 cycles/sec. with a maximum f_q of 20,000 cycles/sec, Eq. (13-43) shows that m_p must not exceed 5 (assuming that E_a has the same maximum amplitude at all frequencies).

With frequency modulation Eq. (13-41) shows that if $kf_0 \gg f_q$ we may write approximately

$$bw = 2kf_0 \tag{13-46}$$

But k is proportional to the amplitude of the modulating signal or

$$k = k_fE_a \tag{13-47}$$

where k_f is a proportionality factor and E_a is again the amplitude of the modulating signal. Substituting Eq. (13-47) into (13-46) gives

$$bw = 2k_fE_af_0 \tag{13-48}$$

which shows that the band width is *constant* for a given amplitude since the carrier frequency f_0 does not vary. Substituting Eq. (13-47) into (13-40) and solving for m_f gives

$$m_f = \frac{k_f\ E_a\ f_0}{f_q} \tag{13-49}$$

which shows that with constant E_a, m_f varies inversely with the frequency of the modulating signal, so that, if a radio transmitter capable of transmitting a band width of 200,000 cycles/sec is modulated with a maximum f_q of 20,000 cycles/sec, the maximum permissible m_f at 20,000 cycles/sec is 5 but at 2000 cycles/sec m_f will be 50 for the same value of E_a, at 200 cycles/sec it will be 500 and at 20 cycles/sec it will be 5000. Thus very large values of m may be obtained at the lower modulating frequencies without

exceeding the band width of the equipment, greatly reducing the effect of noise and other interfering signals at these frequencies.

The foregoing may be summarized by stating that the band width is very nearly proportional to mf_r in either phase or frequency modulation, but with frequency modulation m increases as f_s decreases thus permitting full use of the available band width of the transmitter (or other equipment), whereas with phase modulation m is the same at all frequencies (for a given amplitude E_s) and the full available band width of the radio equipment is not used except at the highest modulating frequencies. Therefore, what would otherwise be phase modulation in the Armstrong transmitter is effectively changed to frequency modulation by inserting a filter in series with the modulating source, as indicated in Fig. 13-29, the output of which is inversely proportional to the impressed frequency. This makes m_p vary inversely as f_s.

The third objection is overcome by the use of frequency doublers Each doubler doubles not only the carrier frequency but also the frequency deviation (the signal at B, Fig. 13-29, may now be considered as a f-m wave due to the presence of the filter in the modulating circuit). Thus the number of doublers is given by

$$2^n = \frac{(kf_0)_n}{(kf_0)_1} \tag{13-50}$$

where n = number of doublers

$(kf_0)_n$ = desired frequency deviation at point F, Fig. 13-29

$(kf_0)_1$ = initial frequency deviation at point B, Fig. 13-29

To obtain an idea of the number of doublers required, recall that the maximum phase shift occurs at the lowest modulating frequency to be transmitted and must not exceed about 0.5 radian. If the lowest modulating frequency is 40 cycles/sec, then $(kf_0)_1 = f_s m_f = 40 \times 0.5 = 20$. Assuming that a band width of 200,000 cycles/sec is desired, Eq. (13-46) shows that $(kf_0)_n = 100{,}000$, so that $2^n = 100{,}000/20 = 5000$. Solving for n shows that slightly over 12 doubler stages are needed.

If 12 doubler stages are used and the final frequency is to be, say, 42 Mc, the original carrier frequency must be $42{,}000{,}000/2^{12} =$ 10,200 cycles/sec. This is much too low a carrier frequency to be successfully modulated for high-quality telephone transmission; consequently an initial carrier frequency of perhaps 200 kc is used.

If this is operated on by a series of six doublers, the resulting frequency at point C, Fig. 13-29, is 12,800 kc, with a minimum frequency deviation of $20 \times 2^6 = 1280$ cycles/sec. At this point the output of another crystal oscillator is introduced, and the resulting signal demodulated by a heterodyne detector.[1] The output will then contain the sum and difference of the crystal frequency and the 12,800-kc carrier, *each carrying the frequency modulation represented by a frequency deviation of* 1280 cycles/sec. The difference frequency should be adjusted to the final frequency divided by the doubling action of the remaining six doublers, or $42{,}000/2^6 = 656$ kc. Therefore, the crystal oscillator at point D of Fig. 13-29 should give a frequency of $12{,}800 - 656 = 12{,}144$ kc (or $12{,}800 + 656 = 13{,}456$ kc). The remaining six doublers will then increase the carrier frequency from 656 kc to the final 42 Mc and will increase the frequency deviation from 1280 cycles/sec to $1280 \times 2^6 = 82{,}000$ cycles/sec.[2]

This method of securing frequency modulation evidently uses a good many tubes owing to the large number of doubler stages. However, the power output required of each doubler is small, and receiving-type tubes may be used. Thus neither the initial investment nor the operating cost is excessive.

Reactance-tube Modulator. Another common method of producing frequency modulation is through the use of a reactance tube. To understand the principles of its operation, first consider the general equation for the alternating plate current of a tube, obtained from Eq. (3-23), or

$$\mathbf{I}_p = \frac{-\mu \mathbf{E}_g}{r_p + \mathbf{Z}_L} \qquad (13\text{-}51)$$

If r_p is very large compared to $\mathbf{Z}_L$, Eq. (13-51) reduces to approximately

$$\mathbf{I}_p = \frac{-\mu \mathbf{E}_g}{r_p} = -g_m \mathbf{E}_g \qquad (13\text{-}52)$$

The conditions of Eq. (13-52) may be most readily realized by using a pentode tube, because of its very high plate resistance.

[1] See p. 578 for a discussion of heterodyne detectors.

[2] This is somewhat less than the 100,000 cycles/sec originally called for, but one more doubler would increase the deviation to 164,000 cycles/sec. Furthermore a slight increase in the original phase shift from 0.5 to 0.61 radian will give a final deviation of 100,000 cycles/sec.

The plate current in such a tube is therefore virtually in phase with the grid voltage regardless of the phase of the plate voltage, since there is no quadrature term in Eq. (13-52). If, then, a plate voltage is applied that is 90 deg out of phase with that of the grid, the voltage and current in the plate circuit will differ in phase by 90 deg and the tube will appear as a reactance to any external circuit coupled to the plate (where the reactance $X = E_p / -I_p$). Furthermore, the equivalent reactance may be varied by altering the grid voltage, since this will vary g_m and therefore I_p. Such is the principle of the reactance-tube modulator of Fig. 13-30.

The plate circuit of the modulator tube (tube 1 of Fig. 13-30) is bridged across the tuned grid circuit of an oscillator (tube 2).[1]

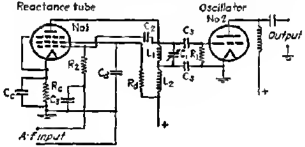


FIG. 13-30. Reactance-tube frequency modulator.

The control grid of the modulator (the one nearest the cathode) is supplied with a voltage of the same frequency as that impressed on the plate through the series circuit consisting of C_2 and R_2. The resistance of R_2 is very much less than the reactance of C_2, so that the current flowing through R_2, and therefore the voltage applied to the grid, differ from the voltage across the tuned grid circuit of the oscillator by practically 90 deg. (The reactance of the condensers C_3 is practically zero at the oscillator frequency.) Application of a-f modulating voltage to the grid through the circuit indicated varies the g_m of the tube and therefore the equivalent reactance $E_p / -I_p$, thus producing frequency modulation of the

[1] Although the plate circuit of this oscillator is not tuned, it contains sufficient inductive reactance to make the tube oscillate in the manner of the tuned-grid-tuned-plate oscillator.

oscillator. The function of the other circuit constants is largely evident from the diagram. C_c, C_d, and C_3 are by-pass condensers, C_3 by-passing the radio frequency only. L_2 is a r-f choke.

The reactance-tube method of modulation has the inherent defect of failing to provide suitable frequency stabilization, since it is impractical to frequency-modulate a crystal oscillator. This problem is met by the addition of a separate frequency-stabilization system controlled by a crystal oscillator, illustrated by the block diagram of Fig. 13-31. The output of a crystal oscillator, the frequency of which differs from that of the transmitter by a low radio frequency,[1] is combined with a portion of the trans-

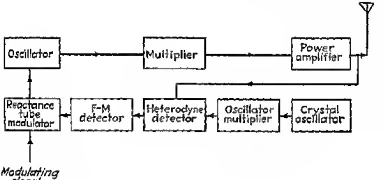


FIG. 13-31. Block diagram of a complete transmitter using a reactance-tube frequency modulator.

mitter output in a heterodyne detector, the output of which is responsive to the difference frequency. A delay circuit is introduced which prevents the output of the heterodyne detector from responding to the rapid variations in frequency produced by the original frequency modulation of the transmitter, and thus the output frequency varies only because of drift in the average frequency of the transmitter. This output is then applied to a f-m detector[2] which produces variations in its direct plate current

[1] The frequency of the transmitter is generally so high (commonly of the order of 100 Mc) as to preclude the use of a crystal operating at the frequency of transmission. A crystal of lower frequency is then used, and its frequency multiplied by suitable doublers or triplers as shown until it differs from the transmitter frequency by the desired amount.

[2] See p. 595 for a discussion of f-m detectors.

proportional to the change in frequency impressed on the detector. The output of this detector alters the average grid voltage (or bias) of the reactance tube and so controls the average frequency of the modulated oscillator.

The modulated oscillator is normally operated at a frequency somewhat lower than the final output frequency in the interest of higher quality of modulation, since a smaller variation in frequency is then required, and its frequency is then multiplied by doublers or triplers up to the final desired value. This part of the process is similar to the Armstrong method described in the preceding section, but the extent of such multiplication is much less, the oscillator frequency being very much higher than that of the Armstrong circuit.

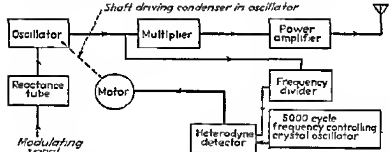


FIG 13-32 General arrangement of a transmitter using synchronized frequency modulation

Synchronized Frequency Modulation. A third system of frequency modulation differs from the foregoing only in the method of securing frequency stabilization. It is known as *synchronized frequency modulation* because of the mechanical synchronizing process that is a basic part of the circuit. The general principle of operation is illustrated in the block diagram of Fig. 13-32.[1]

It may be seen from a comparison of Figs. 13-31 and 13-32 that the main transmitter circuits are essentially the same in these two systems, both utilizing reactance-tube modulation.

[1] A general discussion of this type of transmitter, together with pictures, is given by W. H. Doherty, Synchronized FM Transmitter, *Bell Lab Record*, **19**, p 21, September, 1940, also see paper entitled Synchronized Frequency Modulation, *Communications*, **20**, p 12, August, 1940.

In the circuit of Fig. 13-32, however, the frequency of the oscillator is reduced through a frequency divider until it is of the order of 5000 cycles and is then combined with the output of a crystal oscillator of such frequency that, when the main transmitter is operating normally, the difference in frequency produced in the heterodyne detector is zero. Any change in the transmitter frequency will then produce a difference frequency in the output of the heterodyne detector. The output of the detector is applied to a small motor, mechanically connected to the shaft of the oscillator tuning condenser as indicated, which rotates at a speed proportional to the applied difference frequency. Rotation of the tuning condenser shifts the frequency of the modulated oscillator until it is an exact multiple of the frequency-controlling oscillator, where it is maintained.

The frequency divider consists essentially of a series of heterodyne detectors[1] in each of which one of the two input frequencies is that of the output from the preceding detector (or, in the case of the first detector, the frequency of the modulated oscillator), and the other is obtained from the output of the detector under consideration. Thus the difference frequency appearing in the output of each heterodyne detector is given by

$$f_{out} = f_{in} - f_{out} \tag{13-53}$$

from which it is evident that the output frequency will be half the input frequency.

3. PULSE MODULATION

Intelligence may be transmitted by sending the r-f carrier wave in a series of pulses of very short duration, of the order of 1 μsec. The amplitude of the modulating signal is then represented by shifting the relative *position* of the pulses as in pulse-time (or pulse-position) modulation, by changing the *width* of the pulses producing pulse-width modulation, or by changing the *height* of the pulses to give pulse-amplitude modulation, Fig. 13-33. In each case the frequency of the modulating signal is given by the rate at which the position, width, or amplitude of the pulse changes.

[1] See p. 578 for discussion of heterodyne detectors. The important feature of these detectors, in the present application, is that they produce an output having a frequency equal to the difference between the frequencies of two impressed signals.

(The pulses are shown much wider than they are in practice, the space between pulses normally being much wider than the pulses.)

Pulse-time modulation is illustrated in Fig. 13-33*b* where the rectangular blocks represent the amplitude of the carrier wave and the uniformly spaced broken lines represent the position that the

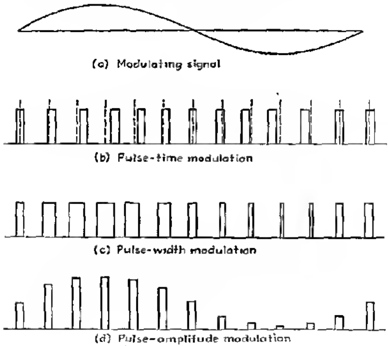


Fig. 13-33. Pulse modulation. Lines (*b*), (*c*), and (*d*) show the amplitude

center of each pulse would occupy if there were no modulating signal. Inspection of the figure will show that the displacement of a pulse from its unmodulated position is proportional to the instantaneous amplitude of the modulating signal given in curve *a*.

The time interval between unmodulated pulses is fixed; thus the number of pulses transmitted per cycle of the modulating signal

decreases as the modulating frequency increases. Experience has indicated that satisfactory transmission of voice can be attained with a pulse repetition rate as low as 2½ times the highest modulating frequency to be transmitted.[1] Thus for the transmission of voice alone (not music) a pulse repetition rate of 8000 times per second is adequate to transmit the required audio range of approximately 200 to 3000 cycles/sec.

Pulse modulation requires a relatively wide band width since a Fourier analysis of the pulse-wave shape will show that rather high frequency components must be included to reproduce each pulse properly. As in frequency and phase modulation, however, the increased band width is accompanied by a marked improvement in signal-to-noise ratio.

An important advantage of pulse modulation is its ready adaptation to multiplex operation. With a pulse repetition rate of 8000 cycles/sec, the time interval between the broken lines of Fig. 13-33*b* is 125 μsec. The pulse width can be of the order of 1 μsec, and a shift of ±5 μsec is usually sufficient to carry the modulation. Thus not more than about 10 μsec out of every 125 is actually required to transmit the intelligence, and the remaining time may be used to transmit other pulse-modulated signals.[2]

There are a number of ways of producing pulse modulation. One is to use a bank of multivibrators, one for each channel of a multiplex system, synchronized by a controlling oscillator. The pulses produced by a given multivibrator are then altered in width by the modulating signal, and the trailing edge of each is differentiated by a suitable circuit to give another pulse which is constant in width but which varies in position with the original modulation.

A more ingenious method and one which combines the multiplexing and modulation processes involves the use of a tube known as the *cyclophon*.[3] This tube consists of an electron gun similar to

[1] D. D. Grieg and A. M. Levine, Pulse-time-modulated Multiplex Radio Relay System-terminal Equipment, *Elec. Commun.*, 23, pp. 159–178, June, 1946.

[2] Equipment has been designed which will transmit 24 separate channels and a synchronizing signal, with a pulse repetition rate of 8000, a pulse width of about 0.5 μsec and a deviation of only ± 2.5 μsec. See Grieg and Levine, *loc. cit.*

[3] See D. D. Grieg, J. J. Glauber, and S. Moskowitz, The Cyclophon: A Multipurpose Electronic Commutator Tube, *Proc. IRE*, 35, pp. 1251–1257, November, 1947.

those used in cathode-ray tubes, electrostatic deflecting plates, a stopper or aperture plate, and a series of targets located opposite the openings in the aperture plate, Fig. 13-34. The electron beam is accelerated and focused in the usual manner by the grid A and the two anodes B and C and is then caused to rotate by applying voltages to the two pairs of deflecting plates, D and E, which differ in phase by 90 deg. This causes the electron beam to rotate over the series of apertures in the stopper plate H thus providing a pulse of current to each target (F_1, F_2, . . .) in succession. The stopper plate is normally maintained at a more positive potential than the targets, thereby producing current flow from targets to stopper plate owing to secondary emission from the latter and

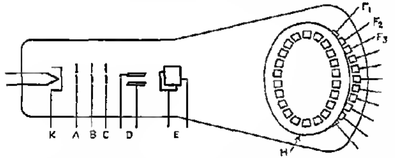


FIG. 13-34. Cyclophon tube for multiplex pulse modulation system

considerably increasing the current flow through the external circuits.

The cyclophon thus produces a series of pulses which are equally spaced, the pulses due to any one target constituting a single channel for the transmission of intelligence. Modulation of these pulses is accomplished by the use of gate clippers, one in series with each target and controlled by the modulation for that channel, which isolate a small slice of the pulse, the width of the slice being proportional to the instantaneous amplitude of the modulating signal. The resulting signal may be used to pulse-width-modulate the carrier wave or it may be translated to pulse-time modulation by using a differentiating circuit to produce a pulse corresponding in time to the trailing edge of the pulse produced by the gate clipper. This latter process will produce an output for each channel exactly

like Fig. 13-33*b*, except that the ratio of pulse width to the space between pulses is much smaller than indicated. Pulses of the other channels from the cyclophon will occupy the spaces between the pulses shown in Fig. 13-33*b*.

A similar cyclophon tube may be used at the receiving end to separate the various channels (see page 599), synchronization being maintained between the two tubes by means of synchronizing pulses. The synchronizing pulses are produced on one of the channels of the cyclophon but differ slightly in shape from the signal pulses so that they may be distinguished at the receiving end.

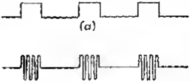


FIG. 13-35. Curve (*b*) shows the actual r-f wave represented by the pulses of (*a*).

The actual transmitted wave will appear somewhat as in Fig. 13-35 where (*a*) shows three of the pulses produced by the cyclophon tube and gate clippers. These are in turn used to key the transmitter so that the radiated wave is a series of sine waves (*b*) of radio frequency. Actually there are many more cycles of radio frequency per pulse than can be depicted; a 1-μsec pulse, for example, is long enough to include 100 cycles of a carrier wave having a frequency of 100 Mc.

PROBLEMS

13-1. A carrier wave of 1000 watts is 30 per cent amplitude modulated. What is the total power radiated?

13-2. A certain class C amplifier with 2000 volts plate supply delivers 75 watts at 56.4 per cent efficiency. It is to be modulated by a class A amplifier using the circuit of Fig. 13-9. Assume that the class C amplifier presents a load impedance of 30,000 ohms to the modulator. The tubes to be used in the modulator are rated at 2000 volts plate supply with $\mu = 19$, $r_p = 3000$ ohms. (*a*) How many of these modulator tubes must be operated in parallel to secure 100 per cent modulation of the class C amplifier with a 2000-volt plate supply if the transformer turns ratio is such as to cause the modulator tube (or tubes) to work into a load resistance of $2r_p$? (*b*) What are the turns ratio and a-c power rating of the transformer? (Maximum permissible excitation on the modulator tube is 48 volts rms.)

13-3. A triode with characteristics as given in Figs. 3-15 and 3-16 is used as a square-law modulator. If $E_b = 90$ volts, $E_c = -5$ volts, carrier voltage $= 1$ rms, modulating voltage $= 0.7$ rms, $\mu = 13.5$, and $R_L = 50{,}000$ ohms, solve for m.

HINT: Graphically solve for r_p and r_p' from Fig. 3-15 or 3-16.

13-4. Rewrite Eq. (13-36) putting in numerical values for all amplitudes and angular frequencies for $m_f = 4$, $f_{carrier} = 50$ Mc, $f_s = 1200$ cycles/sec. Include all side frequencies up to and including $(m_f + 1)$.

13-5. In a certain f-m transmitter $f_{carrier} = 40$ Mc, $bw = 200$ kc. What is the maximum phase shift *either* side of carrier for (*a*) $f_s = 5000$, (*b*) $f_s = 25$.

13-6. Two broadcast signals having frequencies of 1000 and 950 kc are impressed on the grid of a tube which is producing cross modulation by operating on a curved portion of its characteristic curve. (*a*) If only first- and second-order terms are considered, what frequencies are present in the plate current of the tube? (*b*) If third-order terms are also considered, what additional frequencies are present?

CHAPTER 14

DEMODULATORS

A modulated wave was shown in the preceding chapter to be one in which a given signal, usually of low frequency, is superimposed on a carrier wave of higher frequency. The carrier wave is then commonly transmitted by radio or wire to some remote point where the original signal must be removed from the carrier before it can be used. The process by which this is accomplished is known as *demodulation* or, especially with carrier waves of radio frequency, *detection*, and the vacuum tubes and associated circuits used in this process are known as *demodulators* or *detectors*.

Demodulation. Demodulation is defined by the Institute of Radio Engineers in their Standards Report as "the process of modulation carried out in such a manner as to recover the original signal." As implied in this definition, the processes of modulation and demodulation are very similar, the difference, in the square-law type at least, lying primarily in the type of output circuit selected to utilize the products obtained.

1. DEMODULATION OF AMPLITUDE-MODULATED WAVES

In general there are two types of amplitude demodulators: (1) linear (large-signal) detectors and (2) square-law (small-signal) detectors. Linear detectors, when operating ideally, reproduce the original modulating signal without distortion and are, therefore, usually preferred to square-law detectors which introduce some amplitude distortion, usually in the form of a second harmonic of the original modulating signal. The circuits used for the two types are often exactly the same, the difference in performance being due to the amplitude of the applied signal It is for this reason that they are commonly known as *large-signal* and *small-signal* detectors, respectively.

Large-signal Diode Detectors. Figure 14-1 shows the circuit of a simple diode detector. Reference to the material in Chap. 7 will disclose that this circuit is the same as that of a half-wave,

single-phase rectifier with a condenser filter. Thus the output will be pure direct current (except for a small ripple) when a constant alternating voltage is applied, as when the input consists of an unmodulated carrier. An a-m wave, on the other hand, impresses a signal of variable amplitude, and the output of the detector should vary in direct proportion to this changing amplitude. Whether or not it does so evidently depends in large part on the time constant of the RC circuit in the output of the detector as well as on the tube characteristics.

The plate current in the circuit of Fig. 14-1 will flow in a series of short pulses at the peaks of the impressed wave in a manner similar to the performance of a rectifier with condenser-input filter. Normally the resistance R_L is very large so that the current pulses are much shorter than in the usual rectifier circuits and the

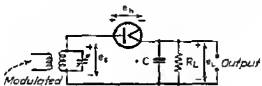


FIG. 14-1. Circuit of a diode detector.

condenser C is, therefore, charged to a potential almost exactly equal to the peak value of the impressed voltage (see discussion of Fig. 7-37, page 199) This type of demodulator is, therefore, frequently known as a *peak linear detector*.

The action of this circuit is pictured in Fig. 14-2 for an impressed modulated wave. Actually the number of r-f cycles per cycle of modulation signal is many times that shown, the actual number being so great as to make an accurate illustration impossible The potential of the cathode is used as reference point, and the voltage across the condenser is indicated by the distance between the dotted line and the Y axis, with the impressed r-f voltage superimposed on the curve of condenser voltage. Since the voltage e_b across the tube is, from Fig. 14-1,

$$e_b = e_s - e_L \qquad (14\text{-}1)$$

and since the condenser voltage, e_L, is here plotted negatively, the voltage across the tube is evidently represented by the total in-

stantaneous distance between the solid curve and the Y axis. Current will, of course, flow through the tube only during the short intervals that this tube voltage is positive, *i.e.*, when the solid curve lies to the right of the Y axis.

The current flow is indicated in Fig. 14-2 with time units corresponding to those of the impressed voltage.

The curves of Fig. 14-2 indicate that the condenser voltage will equal the crest value of the impressed signal and will therefore vary in the same manner as the envelope,[1] *if the tube drop is*

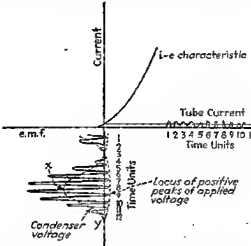


FIG. 14-2. Characteristic curve and alternating voltage and current in a peak linear (large-signal) detector.

negligible, since the locus of the positive peaks of applied voltage would then coincide with the Y axis. Since the tube characteristic must have an infinite slope for the tube drop to be zero, it is evident that reasonably distortionless operation requires that the static characteristic curve have a steep slope.

Large-signal Triode Detectors. A triode may be used as a detector instead of a diode, the circuit being shown in Fig. 14-3 where the tube is biased to cutoff. The general performance of this

[1] Unless the time constant of the R_LC circuit is made too large compared to the time of an a-f cycle (see discussion on p. 571).

circuit is similar to that of the diode; in fact, the tube may be thought of as an amplifier producing a voltage $-\mu e_s$ which is applied to a diode consisting of the plate-cathode circuit of the tube. With this assumption the i-e curve of Fig. 14-2 may represent the $i_b - e_b$ curve of the triode of Fig. 14-3, and the impressed

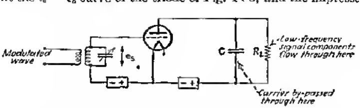


Fig. 14-3. Circuit of a bias-type triode detector

alternating voltage e_s of Fig. 14-2 will be $-\mu e_s$ for the circuit of Fig. 14-3.

Analysis of the Peak Linear Detector. Since the load impedance in the peak linear (large-signal) detector is a function of frequency, analysis by means of Eq. (12-5) is a possible approach, with evaluation of the constants from Eqs. (12-46) and (12-47).

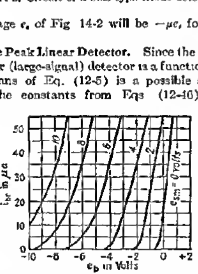


Fig. 14-4. Rectification characteristic curves of a diode

However Ballantine has developed another method in which the performance of this detector is compared with that of an amplifier, which somewhat simplifies the problem.[1]

In this method a set of characteristic curves (Fig. 14-4) is

[1] Stuart Ballantine, Detection at High Signal Voltages, *Proc. IRE*, **17**, p. 1153, July, 1929.

plotted for the diode, similar to the triode static characteristic curves of Fig. 3-16 (page 53), except that the crest value of the impressed r-f signal voltage e_{sm} is one of the independent variables instead of e_c. The plate current i_{br} is the direct, or rectified, component of the plate current due to the presence of the r-f

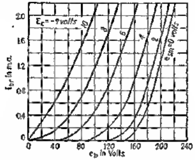


FIG. 14-5. Transrectification characteristic curves of a triode.

signal E_s, the subscript r distinguishing this current from the total plate current i_b which must include r-f components as well. If a triode is used, the same curves are plotted (Fig. 14-5), although a different set must be obtained for each value of direct grid voltage to be used. The curves of Fig. 14-4 are known as *rectification characteristic curves*, and those of Fig. 14-5 as *transrectification characteristic curves*.

The curves of Figs. 14-4 and 14-5 may be determined experimentally, as were the static curves of Fig. 3-16, using the circuit of Fig. 14-6 for a diode and that of Fig. 14-7 for a triode or multigrid tube. An a-f source may be used as E_s (rather than r-f) provided its frequency is not so low that the d-c plate meter will respond to

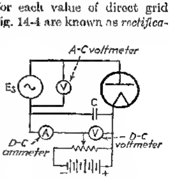


FIG. 14-6. Circuit for determining the curves of Fig. 14-4.

the alternating components in the plate circuit. The condenser C must be of such size that negligible voltage appears across it at the frequency of E_s. For triodes and multigrid tubes the grid-bias voltages at which the curves are determined should be such as to

produce cutoff at the plate and screen potentials available for use with the tube.[1]

The similarity of the curves of Figs. 14-4 and 14-5 (especially the latter) to those of Fig. 3-16 is quite striking, and it will be shown that these curves may be used to determine the performance of

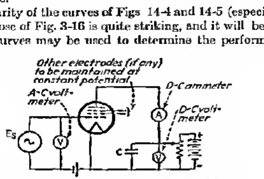


FIG. 14-7. Circuit for determining the curves of Fig. 14-5.

this detector in the same manner as those of Fig. 3-16 were used to determine the performance of a class A amplifier, either by the equivalent circuit of Fig. 3-24 or by the use of load lines as on page 328. For example, if an a-m wave is impressed, the crest value is given by

$$e_{cm} = E_{0m} + mE_{0m} \sin qt \tag{14-2}$$

The equation of the voltage applied to the grid of a class A amplifier is

$$e_c = E_c + E_{gm} \sin qt \tag{14-3}$$

Since comparison of the curves of Figs. 14-4 and 14-5 with 3-16 shows that e_{cm} may be treated as the equivalent of e_c, it is evident that the carrier voltage E_{0m} must be the equivalent of the bias voltage of an amplifier, E_c, and the voltage mE_{0m} the equivalent of the alternating excitation voltage E_{gm}.

The curves of Fig. 14-4 or 14-5 may be expressed mathematically as

$$i_{br} = f(e_{cm}, e_b) \tag{14-4}$$

Equation (14-4) is exactly like Eq. (12-1), page 488, and may therefore be expanded by Taylor's series to give an equation of the

[1] Ballantine has shown that a bias more negative than cutoff will result in improved performance *if the modulation factor m is at all times appreciably less than 1*, see Ballantine, *loc. cit.*

same form as Eq. (12-5). If we write $i_{br} = I_{b0r} - i_{p0r}$ by analogy with the footnote on page 488, we may also write by analogy with Eq. (12-5)

$$-i_{p0r} = d_1 e_a + d_2 e_a^2 + d_3 e_a^3 \cdots \tag{14-5}^*$$

where $e_a = mE_{0m} \sin qt$ and corresponds to e_g in Eq. (12-5). The constants in this equation may be evaluated in the same manner as those for Eq. (12-5) and will be

$$d_1 = -\frac{\dfrac{\partial i_{br}}{\partial e_{sm}}}{1 + R_L \dfrac{\partial i_{br}}{\partial e_b}} \tag{14-6}$$

$$d_2 = -\frac{R_L^2 \left(\dfrac{\partial i_{br}}{\partial e_{sm}}\right)^2 \dfrac{\partial^2 i_{br}}{\partial e_b^2}}{2\left(1 + R_L \dfrac{\partial i_{br}}{\partial e_b}\right)^3} + \frac{R_L \dfrac{\partial i_{br}}{\partial e_{sm}} \dfrac{\partial^2 i_{br}}{\partial e_{sm} \partial e_b}}{\left(1 + R_L \dfrac{\partial i_{br}}{\partial e_b}\right)^2} - \frac{\dfrac{\partial^2 i_{br}}{\partial e_{sm}^2}}{2\left(1 + R_L \dfrac{\partial i_{br}}{\partial e_b}\right)} \tag{14-7}$$

As in the case of the constants a_1 and a_2 of Eqs. (12-6) and (12-7) these derivatives must be evaluated at the operating point. In this case the operating point is determined by the no-modulation direct plate voltage E_b and the carrier voltage E_{0m}.

The term $\partial i_{br}/\partial e_b$ is evidently similar to the plate conductance of the tube except that it must be evaluated in the presence of the normal carrier voltage. The reciprocal resistance $\partial e_b/\partial i_{br}$ is usually preferred and is known as the *detection plate resistance* symbolized in this book by r_{pr}.

The term $\partial i_{br}/\partial e_{sm}$ is of the nature of a transconductance; but since it is the slope of the rectified-plate-current–signal-voltage curve and not the plate-current–grid-voltage curve, it is not the transconductance of the tube. Ballantine has called this term the *rectification factor* in a diode or the *transrectification factor* in a triode, and in this book both are symbolized by g_{mr}.

As in the case of a_1 and a_2 (Eqs. 12-6 and 12-7), d_1 and d_2 may be simplified to a form similar to that of Eqs. (12-46) and (12-47)

$$d_{1(m)} = -\frac{g_{mr} r_{pr}}{|r_{pr} + Z_{L(m)}|} \tag{14-8}^*$$

$$d_{2(m\pm n)} = \frac{g_{mr}^2 r_{pr}^3 r'_{pr}}{2|(r_{pr} + Z_{L(m)})(r_{pr} + Z_{L(n)})(r_{pr} + Z_{L(m\pm n)})|} \tag{14-9}^*$$

Both the detection plate resistance and the rectification or transrectification factor may be determined by measuring the slope of the appropriate characteristic curve. The detection plate resistance may also be measured in a suitable bridge circuit.[1]

It is now possible to apply the entire analysis of the class A amplifier to the study of the peak linear detector. It is necessary only to replace the terms employed in the study of class A amplifiers with the equivalent terms for the detector. Table 14-1 summarizes these substitutions, based on the material of this section.

TABLE 14-1.—SUBSTITUTIONS TO BE MADE IN APPLYING THE ANALYSIS OF CLASS A AMPLIFIERS TO PEAK LINEAR DETECTORS

Symbol for Class A Amplifiers	Corresponding Symbol for Peak Linear Detectors
e_g	e_{gm}
E_s	E_{0m}
E_{gm}	mE_{0m}
i_p	i_{br}
i_b	e_b
i_{p0}	i_{p0r}
r_p	r_{pr}
g_m	g_{mr}

It is important to note that under normal conditions the current i_{br} flows through the resistance R_L of the detector circuit (Fig. 14-1, or 14-3), the impedance of the condenser C being too high at the frequencies present in i_{br} to pass appreciable current, so this resistance corresponds to the load resistance of the equivalent amplifier. The condenser merely by-passes the r-f components that were not included in the current i_{br}. This brings the further conclusion that the relative magnitudes of R_L and C in the detector output circuit must be such as to present a load impedance of R_L to i_{br} and zero impedance to the r-f components.

Distortion in Peak Linear Detectors. The performance of a peak linear detector, when a modulated wave is impressed, may now be determined by replacing e_a in Eq. (14-5) by mE_{0m} sin qt and expanding the equation in the same manner as for the amplifier on page 495. By analogy with Eq. (12-32) this will give for the first-, second-, and third-order terms and the increase in the direct component

[1] The reader is referred to the Standards Report of IRE for recommended circuits for use with both diodes and triodes.

$$i_{p0r} = \frac{d_2(mE_{0m})^2}{2} + \left(d_1 mE_{0m} + \frac{3d_3(mE_{0m})^3}{4}\right)\sin qt \\ - \frac{d_2(mE_{0m})^2}{2}\cos 2qt - \frac{d_3(mE_{0m})^3}{4}\sin 3qt \quad (14\text{-}10)$$

For distortionless operation only the sin qt term should appear in the output so that Eq. (14-10) shows that all coefficients other than d_1 (*viz.*, d_2, d_3, . . .) should be zero or, from Eq. (14-9), r'_{pr} should be zero for all values of e_s.[1] This means that the curves of Figs. 14-4 and 14-5 must be straight lines and, to the extent that they are not, distortion will be present.

If the load circuit of the detector is a *pure resistance* and has the same magnitude at all modulating frequencies *including zero*, a load line may be drawn on Fig. 14-4 or 14-5 whereupon the complete study of amplitude distortion presented for audio amplifiers starting on page 328 is at once applicable by making the substitutions of Table 14-1. The analysis is exactly the same whether the detector is a diode or triode, once the curves of Fig. 14-4 or 14-5 have been obtained. Analysis of the performance of multigrid tubes should follow that of pentode amplifiers and is therefore not appreciably different from that of triodes.

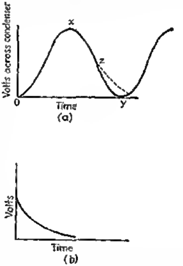


FIG. 14-8. Illustrating the effect of using a load resistance and by-pass condenser with too long a time constant.

Frequency distortion may result from variations in output load impedance with the modulating frequency or from too long a time constant of the condenser and resistance used in the plate circuit of the tube (*e.g.*, C and R_L in the circuit of Fig. 14-1). The latter type of distortion is due to failure of the condenser voltage to follow the envelope at all parts of the modulation cycle. For distortion-

[1] This makes $d_2 = 0$, and if d_2 is zero for all values of voltage, all higher order coefficients are also zero. See footnote on page 495.

less operation it must follow the dashed curve of Fig. 14-2 marked "condenser voltage", but if the time constant R_LC of the detector output is too large, the condenser voltage will deviate from the dashed curve. This is illustrated in Fig. 14-8 where the solid curve in (*a*) represents the dashed curve of Fig. 14-2 drawn to a larger scale, the points x and y corresponding to similarly marked points in Fig. 14-2. From 0 to x the condenser charges as the impressed-signal voltage approaches the positive crest of its modulation cycle; from x to y the impressed voltage decreases, and the condenser must discharge. Since the tube conducts in one direction only, the condenser must discharge through the resistance R_L, the discharge curve being as shown in Fig. 14-8*b*. If the discharge curve is not sufficiently steep (*i.e.*, if the product R_LC is too large), the condenser voltage will follow the dotted curve of Fig. 14-8*a* instead of the solid curve as desired, thereby producing a distorted output wave.

Terman has shown that the time constant of the detector output circuit must satisfy the following relation[1]

$$R_LC \leqq \frac{\sqrt{1 - m^2}}{qm} \tag{14-11}$$

where R_L and C have the significance indicated in Figs. 14-1 and 14-3, and q is 2π times the modulating frequency. Roberts and Williams have extended the foregoing analysis to include the effect of the input impedance to the succeeding tube, giving[2]

$$R_eC \leqq \frac{\sqrt{R_e^2 - m^2R_L^2}}{qmR_L} \tag{14-12}$$

where R_e is the resistance of R_L and the input resistance to the succeeding tube in parallel. Evidently if the input resistance to the succeeding tube is very much larger than R_L, the two preceding equations will give the same results.

The preceding analysis applies equally to diode and triode detectors, but the input impedance of the diode is lower than that of the triode since the diode draws current while the grid of the

[1] F. E. Terman, Some Properties of Grid Leak Power Detectors, *Proc. IRE*, 18, p. 2160, December, 1930.

[2] F. Roberts and F. C. Williams, *Jour. IEE*, p. 379, September, 1934. Also see S. Bennon, Note on Large Signal Diode Detection, *Proc. IRE*, 25, p. 1565, December, 1937.

triode is biased negatively and draws no current. This results in broad tuning and other undesirable effects when a diode is used. The use of a high load resistance (R_L in Fig. 14-1) will greatly reduce this trouble and at the same time will reduce distortion and increase the output voltage. When properly designed, diode detectors are capable of producing low over-all distortion and are more commonly used than are triodes or pentodes.

Another type of distortion, known as *clipping*, results from the almost universal use of automatic volume control in radio receivers, which causes the resistance to the direct component of plate current to be greater than that presented to the alternating components. Clipping will be discussed later in the section on automatic volume control (page 587).

Infinite-impedance Detector. The infinite-impedance detector (Fig. 14-9) eliminates the flow of current that is present in diode detectors and yet retains the high-fidelity characteristics of the

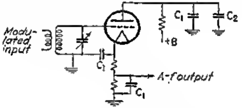


FIG. 14-9. Infinite-impedance detector.

diode. A triode tube is used with the load resistance inserted in the cathode lead, instead of in series with the plate as in the detector circuit of Fig. 14-3. The condensers C_1 have negligible reactance at radio frequency but high reactance at audio frequency, and C_2 has negligible reactance at the modulating frequency. The advantages of this detector may be seen by referring to the curves of Fig. 14-2 which show that the average voltage on a diode (averaged over a r-f cycle) varies with the modulation, as indicated by the dotted condenser-voltage line, this variation being due to the flow of current through the tube at the positive crests of the r-f cycles. In the infinite-impedance detector the average grid voltage varies in a similar manner to the average

voltage on a diode, but in this case the variation is due to the flow of *plate* current which does not flow through the input circuit. Thus the input impedance is high at all times.

It is true, of course, that the triode detector of Fig. 14-3 also has a high-input impedance, but in that type of detector the

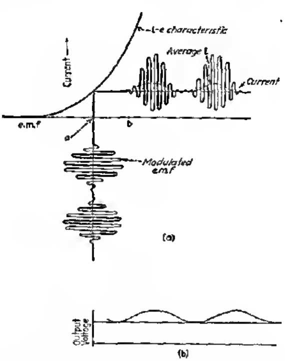


FIG 14-10 Square-law demodulation due to curvature of the characteristic curve of a vacuum tube.

average grid voltage remains constant, and detection is due to variations in the average *plate* voltage. These variations in plate voltage in turn cause the transconductance of the tube to vary throughout the modulation cycle, and some distortion results, i.e., when operating in the circuit of Fig. 14-3 the tube is essentially the equivalent of an r-f amplifier driving a diode detector (the plate circuit).

Square-law (Weak-signal) Detectors. When the applied signal is very weak, the plate current in the detector of Figs. 14-1 and 14-3 may flow throughout the cycle, instead of in pulses as indicated in Fig. 14-2. This produces a performance as indicated in Fig. 14-10. The i-e characteristic curve may represent the current-voltage curve of a diode with zero plate voltage being at a[1], or the plate-current–grid-voltage curve of a triode with zero grid voltage being at b. The 100 per cent modulated wave shown on the vertical axis is applied to the input of the circuit of Fig. 14-1 or 14-3. The approximate shape of the plate current flowing at any instant may then be found with the aid of the i-e characteristic curve and is as shown on the horizontal line of Fig. 14-10.[2] The r-f components of this current are by-passed through the condenser C of Figs. 14-1 and 14-3, and the current flowing through R_L is then as shown in Fig. 14-10b, consisting of a direct component and a component having essentially the same shape as the envelope of the impressed modulated wave.

Performance of a Square-law Detector. The performance of a square-law detector may best be analyzed by means of the power-series expansion. The voltage across R_L may be found by first writing the equation of the incremental plate current, due to application of a signal, from Eq. (12-5) [or from Eq. (12-41)] and then multiplying each component of this current by the impedance presented by the load at the frequency of the component. Let the voltage induced across the resonant circuits of Figs. 14-1 and 14-3 be the modulated wave of Eq. (13-5), page 512. The incremental plate current is then found by inserting Eq. (13-5) into (12-5). Rearranging terms in the order of their frequency and letting $\omega + q = p$ and $\omega - q = r$ in designating the subscripts gives for the first two terms of Eq. (12-5)

[1] It may seem strange that this curve indicates that plate current flows in a diode with zero applied emf. The scale of this figure is such that the difference in potential between point a and the point at which the current drops to zero is probably less than 1 volt, and the magnitude of the current flowing at point a is of the order of a few microamperes. This small current is due to the velocity of emission of the electrons as they leave the cathode and is sufficient for the action of a square-law detector.

[2] This process assumes zero load impedance to all components of the plate current which is, of course, not true, but the results, though not quite accurate, are helpful in visualizing the detector performance as presented in the next section.

$$
\begin{aligned}
i_{p0} = {} & \frac{a_{2(\omega-\omega)}E_{0m}^2}{2} + \frac{a_{2(p-p)}m^2E_{0m}^2}{8} + \frac{a_{2(r-r)}m^2E_{0m}^2}{8} \\
& + \frac{a_{2(\omega-p)}mE_{0m}^2}{2}\sin qt + \frac{a_{2(\omega-r)}mE_{0m}^2}{2}\sin qt \\
& - \frac{a_{2(p-r)}m^2E_{0m}^2}{4}\cos 2qt + \frac{a_{1(r)}mE_{0m}}{2}\cos(\omega-q)t \\
& + a_{1(\omega)}E_{0m}\sin\omega t - \frac{a_{1(p)}mE_{0m}}{2}\cos(\omega+q)t \\
& + \frac{a_{2(r+r)}m^2E_{0m}^2}{8}\cos(2\omega-2q)t + \frac{a_{2(\omega+r)}mE_{0m}^2}{2}\sin(2\omega-q)t \\
& - \frac{a_{2(\omega+\omega)}E_{0m}^2}{2}\cos 2\omega t - \frac{a_{2(p+r)}m^2E_{0m}^2}{4}\cos 2\omega t \\
& - \frac{a_{2(\omega+p)}mE_{0m}^2}{2}\cos(2\omega+q)t \\
& + \frac{a_{2(p+p)}m^2E_{0m}^2}{8}\cos(2\omega+2q)t
\end{aligned}
\qquad (14\text{-}13)
$$

As a first approximation the output impedance of the circuits of Figs. 14-1 and 14-3 may be assumed to be zero at all radio frequencies and equal to R_L at all modulating frequencies and for direct current (ω is a radio frequency and q is a modulating frequency). This means that zero output voltage will be set up by all terms in Eq. (12-13) except the first six. By expanding the a_2 terms by means of Eq. (12-47) and letting $\mu_b' = 0$, $\mu_c' = 0$, the output potential may be written as

$$
e = -\frac{\mu^2 r_p' R_L E_{0m}^2}{4r_p(r_p + R_L)}\left(1 + \frac{m^2}{2}\right) + \frac{\mu^2 r_p' R_L m E_{0m}^2}{2r_p(r_p + R_L)}\sin qt - \frac{\mu^2 r_p' R_L m^2 E_{0m}^2}{8r_p(r_p + R_L)}\cos 2qt \qquad (14\text{-}14)
$$

where $\mu = 1$ for a diode.

Equation (14-14) shows that the output of the detector contains, in addition to an increase in the d-c component, a term of frequency q, which is the desired output, and another of frequency $2q$ which is a second-harmonic distortion component. A comparison of the amplitudes of these last two terms shows that they are in the ratio

$m/4$ so that the percentage of distortion in the output of this detector is

$$\text{Amplitude distortion} = \frac{m}{4} \times 100\%$$

This indicates that the maximum distortion, due to the square-law characteristic of the detector alone, is 25 per cent and that it varies in direct proportion to the percentage of modulation.

Square-law detectors are not widely used, at least partly because of this inherent distortion, except in heterodyne detection (see page 578) where a true square-law detector is theoretically distortionless. However square-law detection commonly takes place in such undesired circumstances as cross modulation (see page 531) and is therefore of more than academic interest.

Components of Square-law Demodulation. A square-law detector functions as a demodulator because it is operated on the curved portion of the tube characteristic curve. Thus it is similar to an amplifier tube producing amplitude distortion and the analysis of the section on page 505 is equally applicable to a detector. Applying this approach permits determination of the frequency of the various components of the plate current without first developing Eq. (14-13). If the characteristic curve is truly square-law so that no terms higher than the second order are present, the plate current of the detector will contain components having the following frequencies: (1) the original frequencies, ω, $\omega + q$, $\omega - q$, (2) the sums of each pair, $\omega + \omega = 2\omega$, $\omega + (\omega + q) = 2\omega + q$, $\omega + (\omega - q) = 2\omega - q$, $(\omega + q) + (\omega + q) = 2\omega + 2q$, $(\omega + q) + (\omega - q) = 2\omega$, $(\omega - q) + (\omega - q) = 2\omega - 2q$ and (3) the differences of each pair, $\omega - \omega = 0$, $\omega - (\omega + q) = q$,[1] $\omega - (\omega - q) = q$, $(\omega + q) - (\omega + q) = 0$, $(\omega + q) - (\omega - q) = 2q$, $(\omega - q) - (\omega - q) = 0$. If the tube characteristic curve is not truly square-law, third-order (and even higher order) terms may appear and produce further components according to the discussion on page 506. The output circuit must be designed to discriminate between these various components and so reproduce only the desired terms, in so far as this is possible.

Grid-leak Detector. Figure 14-11 shows the circuit of the *grid-leak detector*. Essentially it may be considered as a diode detector in which the grid of the triode serves as the anode of an equivalent diode and R_g and C_g serve in the same capacity as C and R_L in Fig. 14-1. Modulation-frequency potentials are developed across

[1] The minus sign is omitted since it has no significance.

R_g in the same manner as across R_L in Fig. 14-1, and these are then applied to the grid of the tube and amplified in the plate circuit. Since the r-f potentials are also applied to the grid, a by-pass condenser must be supplied in the plate circuit as shown to eliminate r-f currents from the output resistance.

The circuit is not widely used because, as may be seen from the curves of Fig 14-10, a sharply curving current characteristic is essential if a large rectified component of current is to be developed in the diode circuit (*i.e.*, the grid circuit), and the grid-current curves of Fig 3-33 show that the greatest curvature is obtainable with zero plate voltage. On the other hand, for maximum gain and minimum distortion in the amplification of the resulting modulation voltage across R_g, the plate voltage should be reasonably high, in direct conflict with the requirements for maximum

FIG 14-11 Circuit of a grid-leak type of triode detector

detecting efficiency in the diode (or grid) circuit. A better procedure is to provide a diode detector and a triode (or better, a pentode) amplifier with separate tubes (although they may be contained in the same envelope) using the diode circuit of Fig. 14-1 followed by a conventional a-f or video amplifier.

Heterodyne Detection. From the discussion of a preceding section it is evident that if two radio frequencies, $\omega/2\pi$ and $q/2\pi$, are simultaneously applied to the input of a square-law detector, the output will contain one term of frequency $(\omega - q)/2\pi = v/2\pi$. If this frequency is very much less than those of the original applied signals, the output circuit may be designed to reject all but this difference frequency. Such a device is known as a *heterodyne or beat detector*.

A simple circuit of a heterodyne detector is shown in Fig. 14-12, using a triode tube. One of the two original signals is indicated as the incoming wave and, in a radio receiver for example, is the signal coming in on the antenna, amplified and then delivered to

the heterodyne detector. The second signal is supplied locally by a vacuum-tube oscillator. The difference frequency is assumed to be low enough to be audible, and the condenser in the output of the detector by-passes all r-f components of the plate current, leaving only the difference frequency, its harmonics due to any dis-

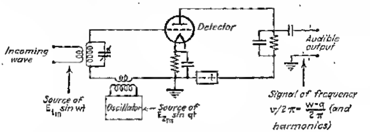


FIG. 14-12. Circuit of a heterodyne detector.

tortion present, and the direct component (which is rejected by the condenser in series with the output terminal).

The circuit of Fig. 14-12 has two principal applications: (1) the reception of unmodulated (continuous-wave) signals and (2) the superheterodyne receiver. Continuous-wave signals consist of a steady carrier wave keyed at stated intervals in accord with a code,

FIG. 14-13. Grid voltage and plate current in a triode detector when receiving radiotelegraph signals (no heterodyne).

usually the International Morse. If such signals are impressed on a conventional detector, the only result is an increase in the direct component of current during the time that one of the signals is being received. Figure 14-13 shows in the upper curve the type of signal received and applied to the grid of the tube. The lower curve shows the plate current flowing in the plate circuit of the detector of Fig. 14-3. If the plate-load resistance happens to be a pair of telephone receivers, the only effect will be to pull the

diaphragm down close to the pole pieces during the interval that the signal is impressed, giving a slight click. The addition of a local oscillator, however, as in Fig. 14-12, will provide an audible tone of any frequency desired in the headphones, the frequency being the difference between that of the incoming wave and that

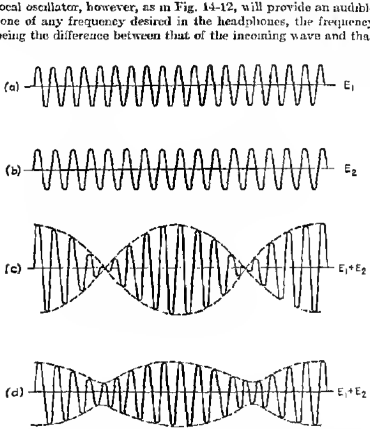


FIG 14-14. Wave shapes in a heterodyne detector circuit

of the local oscillator, the latter usually being under the control of the operator.

The amplitude of the beat-frequency signal in a square-law detector is proportional to the product of the amplitude of the two impressed signals, as shown by the $(\omega - q)$ term of Eq (12-52) The output of the detector may therefore be increased by raising

the amplitude of the locally generated signal, E_2, but if this process is carried too far, the detector will change from square-law to linear operation, whence the amplitude relationships will be somewhat different.

The amplitude relationships in a linear detector may be seen by means of Fig. 14-14. The waves of (*a*) and (*b*) represent the two impressed signals E_1 and E_2 while (*c*) represents their sum when the amplitudes of the impressed signals are equal. The output of a linear detector is proportional to the amplitude of the impressed wave and is therefore proportional to the dotted envelope. A linear detector will, therefore, not produce a pure sine-wave output when the two impressed signals are nearly equal in magnitude but will produce an output having components of frequency $2v/2\pi$, $3v/2\pi$, etc., in addition to the desired term of frequency $v/2\pi$, (where $v = \omega - q$). If, on the other hand, one of the impressed waves is much larger in amplitude than the other, the resulting wave will be as in Fig. 14-14*d*, and the envelope is seen to approach a sine wave as the ratio of the amplitudes of the two impressed waves is increased.

A mathematical treatment of the foregoing may be worked out with the aid of the voltage triangle of Fig. 14-15. Here the incoming signal is represented by the vector $\mathbf{E}_1$ revolving counterclockwise at an angular velocity ω and the local signal by the vector $\mathbf{E}_2$ revolving at a velocity q. The angle between them must then be $(\omega - q)t$ as indicated in the figure. If we now imagine ourselves to be revolving at an angular velocity q, the $\mathbf{E}_2$ vector will appear to stand still and the $\mathbf{E}_1$ vector will revolve at a velocity $(\omega - q)$. Under these circumstances the horizontal and vertical components of the vector representing $(\mathbf{E}_1 + \mathbf{E}_2)$ may be written

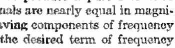


FIG. 14-15. Vector diagram to illustrate the performance of a linear heterodone detector.

$$(\mathbf{E}_1 + \mathbf{E}_2)_h = E_2 + E_1 \cos(\omega - q)t \qquad (14\text{-}15)$$

$$(\mathbf{E}_1 + \mathbf{E}_2)_v = E_1 \sin(\omega - q)t \qquad (14\text{-}16)$$

The actual magnitude of the $(\mathbf{E}_1 + \mathbf{E}_2)$ vector is equal to the square root of the sum of the squares of these two components and therefore varies in a somewhat complex manner with the amplitude

of the incoming wave, but if the amplitude E_2 is very much larger than E_1, the vertical component will have negligible effect and we may use Eq. (14-15) to represent the amplitude of the resulting wave with reasonable accuracy. Since the detector output is proportional to the amplitude of the impressed wave, the E_2 term in Eq. (14-15), being constant in amplitude, will produce a direct component, while the second term will produce a sine-wave component of frequency $(\omega - q)/2\pi = v/2\pi$ in the output of the detector, with an amplitude proportional to E_1. Thus a linear detector will produce a distortionless output if the local signal is very much larger than the incoming signal, but the amplitude of its output will be independent of the amplitude of the local signal, E_2.

Principle of Operation of Superheterodyne Receivers. The superheterodyne receiver makes use of the heterodyne detector to

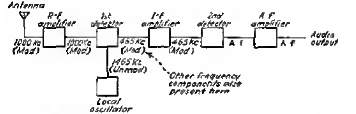


FIG. 14-16. Schematic illustration of the various frequency components appearing in the superheterodyne receiver.

change the incoming signal to one of lower frequency which is then amplified in an intermediate amplifier. The principal advantage is that the intermediate amplifier may be very carefully designed and constructed to give high efficiency in the one frequency band that it is to amplify, whereas the ordinary tuned-radio-frequency (t-r-f) amplifier must be constructed with variable tuning condensers so that it may be tuned to any desired signal throughout a wide range of frequencies. Obviously such an arrangement is much less efficient and less selective. Virtually all modern receivers are of the superheterodyne type.

Figure 14-16 shows a schematic diagram of a superheterodyne receiver. An incoming modulated signal of 1000 kc and an intermediate frequency of 465 kc are assumed. The local oscillator frequency must then differ from 1000 by 465 kc, so it may operate

at either 1465 or 535 kc, preferably the former. The intermediate amplifier is sharply tuned to 465 kc and will therefore reject all other components. It is commonly of a band-pass type (see page 388) to amplify all components of the two side bands equally.

Any modulation carried by the incoming signal is also carried by the i-f[1] signal without distortion if the first detector is truly square-law, or if it is linear and the local oscillator produces a signal that is much larger than the incoming signal. For square-law operation the $(\omega - q)$ term of Eq. (12-52) shows that the detector output is directly proportional to E_1, the amplitude of the incoming signal. If this signal is amplitude-modulated, the effect is as though E_1 varied according to the modulation, and therefore the output

after amplification

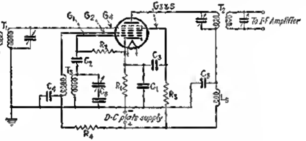


FIG. 14-17. Circuit of a pentagrid convertor tube used as a combination oscillator and first detector in a superhetrodyne receiver.

of the detector also varies in the same manner. Equation (14-15) indicates that a linear detector with a strong local oscillator will perform similarly. The second detector of a superheterodyne receiver is normally a linear detector which recovers the original modulating signal in the usual manner.

It is of interest to note that components at the modulating frequency appear in the plate current of the first detector as well as in the second but, since they are not desired at that point, the output circuit of the first detector is designed to reject them along with the undesired r-f components.

Pentagrid Converter Tubes in Superheterodyne Receivers. In early superheterodyne receivers the circuit used for heterodyning the incoming signal was exactly that of Fig. 14-12, except that the

[1] "I-f" is the commonly used abbreviation for "intermediate-frequency."

output circuit consisted of a tuned circuit resonant to the intermediate frequency, instead of the resistance shown in the figure. Later, mixer tubes were developed which gave superior performance.

A commonly used type of mixer tube is the pentagrid converter, described briefly on page 88. A typical circuit using this tube is shown in Fig. 14-17 where the two inner grids are used as grid and plate of an oscillator, while the other grids and the plate serve as a screen-grid detector for demodulating the combined output of the input signal and the local oscillator. The screen grid, together with proper design of external circuits, eliminates coupling between the local oscillator and the rest of the circuit, except that any electron flow to the plate of the detector unit must flow through the inner grids and so necessarily pulsates at the frequency of the local oscillator. This type of coupling is known as *electron coupling* in that the only connection between the two circuits is through the unidirectional electron stream.

The modulated input is supplied to grid 4 through the air-clad, tuned transformer T_1, grid 4 being biased for detector action by the drop through R_1 (by-passed for radio frequency by C_1). The local oscillations are generated by grids 1 and 2 and the oscillating circuit T_3 and C_5. R_2 and C_2 are the leak and condenser providing grid bias for this oscillator. R_3 drops the direct voltage to a safe value for grid 3, condenser C_3 by-passing the radio frequency to ground. R_4 and C_4 perform the same function for grid 2. The demodulated output is supplied to transformer T_2 which is tuned to the intermediate, or difference, frequency, rejecting all other components.

Pentagrid Mixer Tubes in Superheterodyne Receivers. The principal objection to the pentagrid converter is that the electron stream, pulsating at the frequency of the local oscillations, will cause current to flow not only in the plate circuit but also in the control-grid circuit (grid 4). This becomes increasingly serious as the frequency is increased and the difference between the frequencies of the incoming and local signals decreases. The use of a pentagrid mixer tube (described briefly on page 88) in the circuit of Fig. 14-18 largely eliminates these troubles.[1] The control

[1] See also C. F. Nesslage, E. W. Herold, and W. A. Harris, A New Tube for Use in Superheterodyne Frequency Conversion Systems, *Proc. IRE*, **24**, p. 207, February, 1936.

grid in this tube is the inner one G_1; the local oscillations, generated by a separate tube, are applied to the third grid. Thus the local oscillator cannot affect the control grid; and although the input signal can react on the local oscillator through grid 3, this is of no importance, as the signal strength of the oscillator is nor-

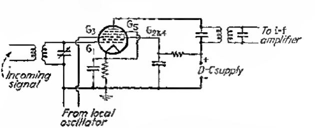


FIG. 14-18. Circuit of a heterodyne detector using a pentagrid mixer tube.

mally many times that of the incoming signal. Grid 1 is of variable-μ construction, and its bias is varied when automatic volume control is used (see page 587). Grid 3 acts essentially in the same manner as the suppressor grid of a suppressor-modulated pentode (Fig. 13-15) and thus produces a difference frequency due to modulation of the electron stream.

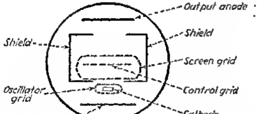


FIG. 14-19. Cross section of the 6K8 type of converter tube.

Another solution to the problem of eliminating the interaction between local oscillator and input circuit is the triode-hexode design, shown in cross section in Fig. 14-19. The oscillator anode is mounted on one side of a flat cathode, and the output anode and control grid are mounted on the other side. The shield shown is grounded internally to the metal envelope of the tube and serves

as a suppressor grid as well as tending to form electron beams. The oscillator grid is seen to surround the cathode completely and therefore modulates the electron stream to the output anode, yet the field of the oscillator anode does not affect the control grid owing to the arrangements of shields.

The performance of both pentagrid converters and pentagrid mixers may be expressed in terms of the *conversion transconductance* of the tube. Conversion transconductance is defined in the Standards Report of IRE as

. . . the quotient of the magnitude of a single beat frequency component $(f_1 + f_2)$ or $(f_1 - f_2)$ of the output-electrode current by the magnitude of the control-electrode voltage of frequency f_1, under the conditions that all electrode voltages and the magnitude of the electrode alternating voltage f_2 remain constant and that no impedances at the frequencies f_1 or f_2 are present in the output circuit. As most precisely used, the term refers to an infinitesimal magnitude of the voltage of frequency f_1.

The performance of these tubes may therefore be predicted from Eq. (10-13) (page 388) by replacing g_m by the conversion transconductance.

Autodyne, or Oscillating Detector. It is quite possible to combine the detector and oscillator of Fig. 14-12 in a single unit, as

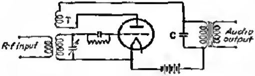


FIG. 14-20. Circuit of an autodyne detector. The same circuit may also be used as a regenerative detector.

in Fig. 14-20. A *tickler* coil T is added in the plate circuit of the detector to feed a small amount of energy from the plate circuit back into the input circuit. The amount of coupling between this tickler and the tuning inductance must be adjusted until oscillations are set up, the frequency being determined by the constants of the input circuit.[1] This circuit can be detuned slightly from the incoming signal without materially reducing the receiver

[1] It should be evident that any circuit that will produce oscillations could be used here, the one shown being simple and easily controlled.

response, until the difference is a suitable audio frequency, say 1000 cycles. If the incoming signal is of frequency $\omega/2\pi$, the tube will be tuned to and will oscillate at a frequency of $\omega/2\pi \pm 1000$, and yet the impedance of the tuned circuit will be sufficiently high at a frequency of $\omega/2\pi$ to give a good response to the incoming signal. This arrangement is not satisfactory for a superheterodyne, where the difference frequency is so great as to require an impossible amount of detuning, but it is eminently satisfactory for the reception of undamped, radio-telegraph signals.

Regenerative Detectors. If the tickler coil in Fig. 14-20 is adjusted to a point where the tube will almost, but not quite, oscillate, any incoming signal will be greatly increased in magnitude over what it would have been without the tickler connection. This action is known as *regeneration* and is due to feeding back sufficient energy from the plate circuit to supply a *portion* of the losses in the input circuit. The incoming signal need then supply only the small remaining losses, which it will do with an increase in amplitude, the net effect being to decrease the effective resistance of the input circuit.

If the coupling of the tickler coil is increased to the point of oscillation, all the losses are supplied by the plate circuit through the tickler connection, and the incoming signal loses full control of the current flowing in the tuned circuit. Reception of radiotelephone signals under these conditions will result in a jumble of sound instead of intelligible speech or music, but a reduction of the feedback to a point just below that required for oscillation will again permit the radiotelephone signals to be clearly heard.

This regenerative feature is one that has proved quite valuable in the operation of small receivers, where a very large increase in sensitivity may be obtained by the addition of a tickler coil or other means of producing regeneration. It is rather critical in adjustment, however, and is therefore seldom used in broadcast or other receivers that are to be operated by untrained hands. In such services the necessary gain is supplied by additional stages of amplification.

Automatic Volume Control. Most modern receivers are equipped with automatic volume control to maintain the output to the loud-speaker essentially constant regardless of the amplitude of the r-f signal impressed on the input to the receiver from the antenna. This automatic action is generally secured by varying

large condenser. The final filter (R_4, C_4) eliminates the direct component, allowing only the a-f components to appear across R_4, which supplies voltage to the first a-f amplifier. The resistance R_4 is commonly in the form of a potentiometer to provide a means of regulating the level of the output signal at the will of the listener.

Circuits of this type may cause clipping of the signal wave due to the difference in the impedance offered by the output circuit to the a-f and direct components.[1] The reason for this may be seen by recalling that the current flowing through a diode is zero when no signal is impressed and that the amplitude of a 100 per cent modulated wave drops to zero at each negative crest of the modulation cycle. Thus the output current, after the radio frequency has been filtered out, will be as shown in Fig. 14-22*a* when the

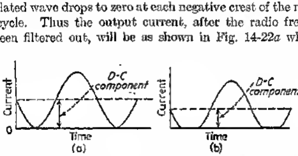


FIG. 14-22. Type of distortion which may take place when the same tube is called upon to supply both automatic-volume-control bias and audio output.

same load impedance is presented to both the direct and the modulating frequency components. With automatic-volume-control circuits the impedance to the direct component may be higher than that presented to the a-f component, resulting in a reduced direct current. This is equivalent to lowering the average (dotted) line in Fig. 14-22*a* which would require current flow through the diode in a negative direction during a small part of the cycle if the a-f wave is not to be distorted. Since this is impossible, the actual output current will appear as in Fig. 14-22*b*, with consequent distortion.

This action may be more fully studied with the aid of the curves of Figs. 14-4 and 14-5. Let us first simplify the circuit of Fig. 14-21 to that of Fig. 14-23 where R_r represents the actual resistance of the output circuit of Fig. 14-21 to the direct component and R_L is the resistance of the a-f output circuit, the condensers being

[1] See Harold A. Wheeler, Design Formulas for Diode Detectors, *Proc. IRE*, 26, p. 745, June, 1938.

assumed of such size as to have negligible effect on the performance of the circuit at the audio-frequencies present. We may then draw a load line for R_v on the curves of Fig. 14-4, as shown in Fig. 14-24, to determine the operating point P for an assumed unmodulated carrier wave of 4 volts crest amplitude. This construction is quite similar to that employed in the graphical study

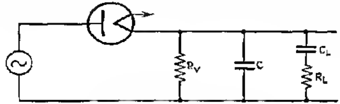


FIG 14-23. Simplified circuit of Fig. 14-21.

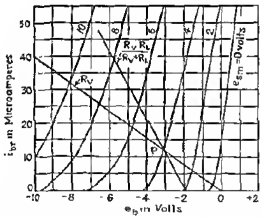


FIG. 14-24. Load lines for the circuit of Fig 14-23 to illustrate how the distortion of Fig 14-22 is generated.

of audio amplifiers on page 322, when the substitutions of Table 14-1, page 570, are made.

When modulation is applied to the impressed carrier, the effect is the same as though the amplitude of the carrier was varied, and it might at first be thought that the performance of the circuit would be indicated by moving up and down the R_v load line as

the amplitude of the impressed signal varies. But the impedance presented to a-c variations in the plate current is lower than that presented to the direct component, being equal to the resistance of R_L and R_v in parallel, and therefore the dynamic performance is found by following the curve marked $R_vR_L/(R_v + R_L)$. This curve, it may be seen, reaches zero before the r-f amplitude has fallen to zero.[1] Thus a 100 per cent modulated wave would have an r-f amplitude varying from zero to 8 volts crest, but the plate current would be zero for all values of impressed r-f voltages less than about 1.9 volts crest, producing the output wave of Fig. 14-22*b*. If the circuit of Fig. 14-1 had been used, the dynamic and static load lines would have been the same and the output would have appeared as in Fig. 14-22*a*.

Distortion due to this cause will not be present if the modulation is sufficiently less than 100 per cent, since the peaks of the a-f component will then be less than the direct component and the total current of Fig. 14-22*b* will never drop to zero at any point in the cycle. Normally the modulation of an incoming wave is constantly varying, and this distortion will therefore be present only during those intervals when modulation exceeds the critical value. Furthermore the input impedance to the detector is affected by the output impedance in such a manner that the impedance to the side bands, in the circuit of Figs. 14-21 and 14-23, is somewhat less than that presented to the carrier. Thus the internal drop in the circuit driving the detector is greater for the side bands than for the carrier, reducing by a small amount the per cent modulation of the signal applied to the detector. This effect appreciably reduces the distortion that would otherwise be present.

Vacuum-tube Voltmeter. The direct component of plate current in a detector of any type has been shown to be a function of the amplitude of the applied potential; therefore, any detector may be used as a voltmeter by inserting a suitable d-c ammeter in the plate circuit. A common method is to use the triode circuit of Fig. 14-3 but with a series condenser added in the input to prevent a change in bias by any direct component present in the potential to be measured, as C_1 in Fig. 14-25. The resistance R in this figure

[1] Some error is involved in this method if the curves of Fig. 14-4 are sufficiently nonlinear to produce an appreciable change in the d-c component as the modulation is applied or removed. See footnote on page 337 which refers to the same phenomenon in class A amplifiers.

is provided to complete the grid circuit for direct current and should be high (several megohms) as a principal advantage of the vacuum-tube voltmeter is that it may be built with a very high input impedance and thus draw negligible power from the circuit under measurement. The condenser C_0 should be large enough to by-pass components of the lowest frequency to be expected.

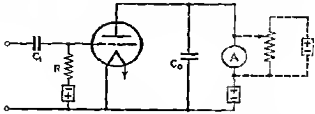


FIG 14-25 Simple type of vacuum-tube voltmeter.

Maximum sensitivity is obtained by biasing the tube to a point near cutoff so that the plate current without impressed signal is only a few microamperes For the measurement of very small voltages, the sensitivity may be increased by balancing out this small zero-signal current, permitting an even more sensitive microammeter to be used. Only the *change* in plate current due to the

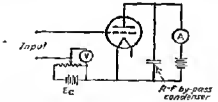


FIG. 14-26 Circuit of a slide-back vacuum-tube peak voltmeter.

impressed signal will then be read on the meter. A simple way of doing this is shown by the dashed lines in Fig 14-25 where the *potentiometer is adjusted until the microammeter reads zero* with no voltage applied at the input to the voltmeter. In actual practice this may be accomplished by a sort of bridge circuit, thus eliminating the need for an additional battery.[1]

[1] As an example, see F E. Terman, "Radio Engineers' Handbook," p. 931, McGraw Hill Book Company, Inc , New York, 1943

Another type of vacuum-tube voltmeter uses a bias-type detector with an adjustable bias (Fig. 14-26). The bias E_c is first adjusted to a, Fig. 14-27a, to give only a few microamperes of plate current when there is no alternating voltage applied. The bias is then increased appreciably, the potential to be measured

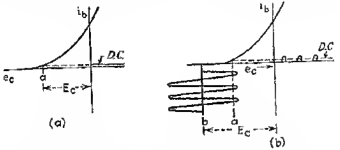


FIG. 14-27. Illustrating the performance of the voltmeter of Fig. 14-26.

is applied to the grid, and the bias is readjusted to b, Fig. 14-27b, to give the original direct plate current. It should be evident from this figure that if the direct plate current is made almost zero, the potential difference between a and b is very nearly equal to the peak value of the impressed emf. The reading of the volt-

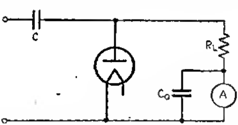


FIG. 14-28. Modification of the circuit of Fig. 14-1 to increase its usefulness as a voltmeter.

meter is obtained by taking the difference between the two voltmeter readings; therefore, the instrument is essentially a peak voltmeter.

The diode circuit of Fig. 14-1 may also be used as a vacuum-tube voltmeter by inserting a microammeter in series with R_L, but it has a disadvantage over the triode in that it must draw

power from the external circuit sufficient to operate the meter and supply the losses in the tube and R_L. By using a very sensitive microammeter the resistance of R_L may be made very large so that the input impedance is many times higher than that of conventional-type voltmeters.

Another disadvantage of the circuit of Fig. 14-1 as a voltmeter is that the circuit under measurement must present a continuous path for the direct component of plate current through the diode This may be avoided by using the modified circuit of Fig. 14-28 where C and R_L serve the same functions as in Fig. 14-1.

The sensitivity of the diode voltmeter may be increased by amplifying its output with a d-c amplifier as indicated in Fig. 14-29. If the amplifier is supplied with a large amount of negative feedback, it will maintain its calibration over long periods of time.

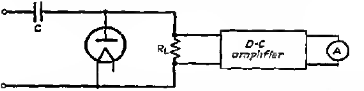


FIG. 14-29. Use of a d-c amplifier to increase the sensitivity of the voltmeter of Fig 14-28.

Characteristics of Vacuum-tube Voltmeters. The response of vacuum-tube voltmeters is not linear with respect to the amplitude of the impressed potential. If the triode of Fig. 14-25 is biased to a point slightly above cutoff, this circuit, as well as the diode circuits of Figs. 14-28 and 14-29, will act as square-law detectors at very low signal inputs, whence the ammeter reading will be nearly proportional to the square of the rms value of the voltage being measured. As the impressed voltage is increased, the response will be more nearly that of a linear detector and the output will be nearly proportional to the peak value of that half of the impressed wave which makes the grid (or anode in the diode) more positive. This means that calibration of vacuum-tube voltmeters is appreciably affected by wave shape. Also vacuum-tube voltmeters are subject to "turnover"; *i.e.*, they may give a different reading if the input terminals are interchanged This is because an impressed wave which contains even harmonics may have a different peak voltage on one half cycle than on the other

The voltmeter circuits presented here will have one terminal grounded if the direct potentials are supplied by a rectifier, as is usually the case. It is important in making measurements that the ground terminal of the voltmeter be connected to the ground terminal of the circuit to be measured.

2. DEMODULATION OF FREQUENCY-MODULATED WAVES

Fundamentally, demodulation of f-m waves is accomplished by first converting variations in frequency into variations in

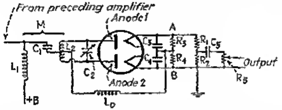


FIG. 14-30. Discriminator circuit used to demodulate f-m signals.

amplitude. This is usually done by applying the f-m wave to a vacuum-tube circuit known as a *discriminator*, in which the output voltage is proportional to the variations in impressed frequency.

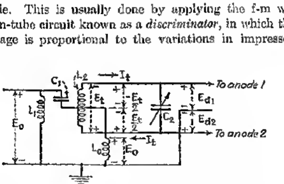


FIG. 14-31. Discriminator circuit showing assumed positive directions of emfs and currents.

A typical discriminator circuit is shown in Fig. 14-30, being an adaptation of the automatic-frequency-control circuit once used in radio receivers.[1] A r-f transformer, consisting of the two coils L_1 and L_2, tuned by the condenser C_2, supplies the two anodes of a

[1] See Charles Travis, Automatic Frequency Control, *Proc. IRE*, 23, p. 1125, October, 1935.

small double diode. The condensers C_1, C_3, and C_4 have negligible reactance at the incoming frequency so that the coil L_0 is effectively in parallel with the primary coil L_1 for radio frequency while providing a return path for the rectified direct current. The voltage built up across the resistances R_3 and R_4 contains direct and modulating components only, since the r-f components are bypassed by C_3 and C_4. The remaining portions of the circuit consist of filters to remove any traces of carrier and to remove the direct component.

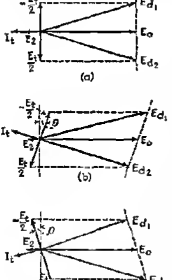


FIG. 14-32. Vector diagrams illustrating the performance of the discriminator circuit of Fig. 14-30 (E_2 is exaggerated for clarity.)

The performance of the circuit of Fig. 14-30 may best be analyzed by redrawing the r-f parts of the circuit as in Fig. 14-31, where the plus and minus signs on the emfs and the arrows for the currents follow the notation of Appendix E. The input voltage is denoted by $\mathbf{E}_0$, and, as previously indicated, essentially this same voltage also appears across L_0, the drop through C_1 being negligible. A voltage $\mathbf{E}_2$ is induced into the secondary L_2 out of phase with $\mathbf{E}_0$ by 180 deg and at resonance will cause an in-phase current $\mathbf{I}_t$ to flow through the coil L_2 and the condenser C_2, as in the vector diagram of Fig. 14-32*a* (where the magnitude of $\mathbf{E}_2$ has been exaggerated). This current flowing through the coil L_2 will produce a reactance drop that is out of phase with the current by virtually 90 deg. In a suitable circuit the resistance is so low (Q so high) that this voltage is many times the induced emf, $\mathbf{E}_2$, therefore, the voltage $\mathbf{E}_t$ across the coil may be considered as being equal to the reactance drop. Since the positive direction of the current $\mathbf{I}_t$ was assumed to be from the bottom to the top of the coil, the voltage appearing between the bottom terminal and the center

tap will be $E_t/2$, whereas that appearing between the top terminal and the center tap will be $-E_t/2$. These voltages are indicated in Fig. 14-31, and the vectors are drawn in Fig. 14-32*a*.

The voltage applied to each diode is evidently the sum of the voltages across the coil L_0 and across the appropriate half of the coil L_2. Therefore, letting E_{d1} and E_{d2} be the voltages between the anodes and the cathode of the tube,

$$E_{d1} = E_0 + \left(-\frac{E_t}{2}\right) \tag{14-17}$$

$$E_{d2} = E_0 + \frac{E_t}{2} \tag{14-18}$$

where the addition must be carried out vectorially as in Fig. 14-32*a*. It is evident from the figure that the *magnitudes* of the voltages E_{d1} and E_{d2} are equal, and therefore the rectified current (which contains the original modulation) is the same in both anodes. The net voltage set up across the two resistances R_3 and R_4 under these conditions is zero, resulting in zero output to the load.

As the frequency varies with the modulation, the voltages applied to the two tubes become unequal, as indicated in Fig. 14-32*b* for a frequency higher than the resonant frequency of the tuned circuit of Fig. 14-30. At this frequency the tuned circuit presents an inductive reactance causing the current I_t to lag E_2 and shifting the vectors $E_t/2$ and $-E_t/2$ as shown. As a result, the voltage impressed on anode 1 is greater than that on anode 2, and the greater rectified current now flowing through anode 1 causes a net voltage to appear across R_3 and R_4 which is positive toward A.

At a frequency below resonance, conditions are as shown in Fig. 14-32*c*, with anode 2 receiving the higher voltage, and the net output voltage across the resistances R_3 and R_4 being positive toward B.

The voltage appearing across these two resistances may be plotted against impressed frequency, giving a curve such as Fig. 14-33. With proper design this curve is very straight over a wide range of frequency variation, giving linear demodulation. It is essential that the straight portion of this curve extend over a frequency band at least equal to the band width of the incoming f-m signal.

Amplitude Limiter. Frequency-modulation receivers must also be equipped with a limiter tube to eliminate any variations in

amplitude before the signal is applied to the detector. This is especially important, as such variations are largely due to noise and other interference, since the originally transmitted f-m signal is constant in amplitude. A typical circuit is shown in Fig. 14-34, where the output coil L_3 is the same coil L_1 shown in the circuit of Fig. 14-30; *i.e.*, the limiter tube is usually placed just ahead of the discriminator. The limiter tube is operated with a grid-leak bias and with rather low plate and screen-grid voltages, obtained by inserting the resistance R_1. Thus grid current flows as soon as signal is applied and produces a bias that increases with the magnitude of the applied signal. The increasing bias decreases the gain of the tube and so maintains a nearly constant output voltage for any signal in excess of a given minimum. The amplitude distortion introduced is not objectionable, since the detector responds to variations in frequency only.

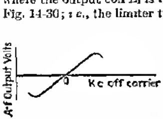


FIG. 14-33. Voltage across points A and B, Fig. 14-30, as a function of the changes in impressed frequency.

3. DEMODULATION OF PULSE-MODULATED WAVES

Demodulation of pulse-width and pulse-amplitude modulated waves (Fig. 13-33*c* and *d*) offers no problem since the average

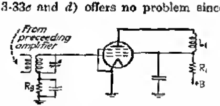


FIG. 14-34. Limiter tube for use with a f-m receiver.

value of these waves, averaged over the pulse-repetition cycle, is proportional to the amplitude of the original modulating signal. Thus an a-m detector, together with a filter having a cutoff frequency slightly below the pulse-repetition frequency, is satisfactory.

Pulse-time modulation requires a somewhat more elaborate system. One method is to superimpose the incoming pulse on a sawtooth wave from a multivibrator, triggering the multivibrator by

the marker pulse to keep it in step with the zero-modulation position of the pulse. This will produce an output wave much as in Fig. 14-35, where the dashed lines indicate the zero-modulation position of the pulses and the multivibrator output is indicated by the saw-tooth waves. The incoming pulses are shown superimposed on the multivibrator output and produce what is effectively a pulse-amplitude type of modulation which may be readily demodulated.

FIG. 14-35. Illustrating the process of converting pulse-time modulation into pulse-amplitude modulation.

With multiplex operation the cyclophon (Fig. 13-34) may be used both to separate the channels at the receiving end and to produce demodulation. The rotation of the electron beam is synchronized with that of the sending-end cyclophon by means of marker pulses which have a different shape from the pulses that carry the modulation. The tube is normally biased to cutoff, but when a pulse is received, the control grid is made less negative and beam current flows, passing through the proper aperture. Since the position of the pulse has been shifted from its zero-signal position by the modulation, only a certain part of the electron beam will pass through the aperture as indicated in Fig. 14-36. In this figure one of the apertures of the cyclophon is indicated by the larger rectangles, and the position of the electron beam, as released by reception of a pulse, is indicated by the smaller, shaded rectangles. Figure 14-36*c* shows the position of the beam relative to the aperture at zero modulation; *a* and *e* show the positions for maximum positive and negative modulation, respectively; and *b* and *d* represent intermediate positions. Only electrons in that

FIG. 14-36. Illustrating the conversion of pulse-time modulation to pulse width modulation by means of the cyclophon tube.

part of the beam which is over the aperture can reach the target and so produce a signal in the output amplifier. From this figure it is evident that the cyclophon tube converts pulse-time modulation to pulse-width modulation after which recovery of the original modulating signal is easy, using conventional circuits

Pulse modulation has thus far found its principal applications in commercial installations where both transmitter and receiver are under the control of trained personnel. The details of forming the pulses, of synchronizing transmitter and receiver, and of other necessary circuit components cannot be covered here and the reader is referred to the technical literature.[1]

PROBLEMS

14·1. The tube to which the characteristic curves of Fig. 14·1 apply is to be operated as a peak linear detector in the circuit of Fig. 14·1 The incoming modulated wave is given by Eq (13-3) in which $E_{am} = 4$ volts, $m = 1.0$, and $q/2\pi$ represents a single modulating frequency. Determine the magnitude of the output voltage at the modulating frequency across a load resistance of 250,000 ohms, and determine the percentage of second harmonic

14·2. A triode with characteristics as given in Figs 3·15 and 3·10 is used as a square-law demodulator $E_b = 90$ volts, $E_c = -5$ volts, $E_{am} = 0.8$ volt, $\mu = 13.5$ If $R_L = 100{,}000$ ohms at audio frequencies and zero at the carrier and higher frequencies, solve for the rms a-f output voltage at the fundamental audio modulation frequency and the percentage of second harmonic for (a) $m = 1.0$, (b) $m = 0.5$, (c) $m = 0.1$.

HINT Graphically solve for r_p and r_p' from Fig 3·15 or 3·10

14·3. A carrier frequency of 1200 kc is amplitude-modulated by a signal of 800 cycles/sec. If this modulated wave is impressed on a truly square-law detector, (a) components of what frequency are to be found in the plate current of the detector? (b) Which of these components will produce an output voltage in the usual detector output circuit?

[1] For further details, see D. D Grieg and A M. Levine, Pulse time-modulated Multiplex Radio Relay System—Terminal Equipment, *Elec. Commun*, 23, pp 159–178, June, 1946, D D Grieg, J. J. Glauber, and S Moskowitz, The Cyclophon A Multipurpose Electronic Commutator Tube, *Proc. IRE*, 35, pp. 1251–1257, November, 1947

APPENDIX A

DEFINITIONS AND NOMENCLATURE

In order to simplify the discussion of vacuum-tube engineering and practice a number of terms and phrases have come into common usage. The student should be thoroughly familiar with these in order to expedite his work. A list of the more commonly used terms and phrases, together with their definitions, is given below. Many of these were taken from the Standards Report of IRE.

A Power Supply. An A power supply is a power-supply device that provides power for heating the cathode of a vacuum tube.

Amplification Factor. Amplification factor is the ratio of the change in plate voltage to a change in control-electrode voltage under the conditions that the plate current remains unchanged and that all other electrode voltages are maintained constant. It is a measure of the effectiveness of the control-electrode voltage relative to that of the plate voltage upon the plate current. The sense is usually taken as positive when the voltages are changed in opposite directions. As most precisely used, the term refers to infinitesimal changes as indicated by Eq. (3-24a). Amplification factor is a special case of μ factor. (See pages 57 and 67.)

Anode. An anode is an electrode to which a principal electron stream flows.

Audio Frequency. An audio frequency is a frequency corresponding to a normally audible wave. The range is roughly from 20 to 15,000 cycles/sec.

B Power Supply. A B power supply is a d-c power-supply device connected in the plate circuit of a vacuum tube.

Blocking (or Stopping) Condenser. A blocking condenser is a condenser used to introduce a comparatively high impedance in some branch of a circuit for the purpose of limiting the flow of l-f alternating or direct current without materially affecting the flow of h-f alternating current.

By-pass Condenser. A by-pass condenser is a condenser used

to provide an a-c path of comparatively low impedance around some circuit element.

C Power Supply. A C power supply is a d-c power-supply device connected in the circuit between the cathode and grid of a vacuum tube so as to supply a grid bias

Carrier. Carrier is a term broadly used to designate carrier wave, carrier current, or carrier voltage.

Carrier Wave. In a frequency-stabilized system, the carrier wave is the sinusoidal component of the useful part of a modulated wave, whose frequency is independent of the modulating wave. (See also pages 510 and 563.)

Cathode. A cathode is the electrode that is the primary source of an electron stream.

Choke Coil. A choke coil is an inductor inserted in a circuit to offer relatively large impedance to alternating currents.

Class A Amplifier.[1] A class A amplifier is an amplifier in which the grid bias and alternating grid voltages are such that plate current in a specific tube flows at all times.

Class AB Amplifier.[1] A class AB amplifier is an amplifier in which the grid bias and alternating grid voltages are such that plate current in a specific tube flows for appreciably more than half but less than the entire electrical cycle.

Class B Amplifier.[1] A class B amplifier is an amplifier in which the grid bias is approximately equal to the cutoff value so that the plate current is approximately zero when no exciting grid voltage is applied and so that plate current in a specific tube flows for approximately one-half of each cycle when an alternating grid voltage is applied.

Class C Amplifier.[1] A class C amplifier is an amplifier in which the grid bias is appreciably greater than the cutoff value so that the plate current in each tube is zero when no alternating grid voltage is applied and so that plate current in a specific tube flows for appreciably less than one-half of each cycle when an alternating grid voltage is applied.

Control Electrode. A control electrode is an electrode upon which a voltage is impressed to vary the current flowing between

[1] To denote that grid current does not flow during any part of the input cycle, the subscript 1 may be added to the letter or letters of the class identification, as A_1. The subscript 2 may be used to denote that grid current flows during part of the cycle.

two or more other electrodes. It is most commonly in the form of a grid.

Control Grid. A control grid is a grid, ordinarily placed between the cathode and an anode, for use as a control electrode.

Control-grid–plate Transconductance. Control-grid–plate transconductance is the name for the plate-current-to-control-grid-voltage transconductance. This is ordinarily the most important transconductance and is commonly known as the *mutual conductance*. (See pages 61 and 67.)

Conversion Transconductance. Conversion transconductance is the quotient of the magnitude of a single-beat frequency component $(f_1 + f_2)$ or $(f_1 - f_2)$ of the output-electrode current by the magnitude of the control-electrode voltage of frequency f_1, under the conditions that all direct electrode voltages and the magnitude of the electrode alternating voltage f_2 remain constant and that no impedances at the frequencies f_1 or f_2 are present in the output circuit. As most precisely used, the term refers to an infinitesimal magnitude of the voltage of frequency f_1. (See page 586.)

Cutoff Grid Voltage. The cutoff grid voltage is that voltage which, when applied to the grid, is just sufficient to reduce the plate current to zero. It is a function of the plate voltage.

Diode. A diode is a two-electrode vacuum tube containing an anode and a cathode.

Excitation Voltage. The excitation voltage is the alternating voltage applied to the input circuit of a tube.

Filament. A filament is a cathode of a thermionic tube, usually in the form of a wire or ribbon, to which heat may be supplied by passing current through it.

Fundamental Frequency. A fundamental frequency is the lowest component of a phenomenon where all the original components are present.

Gas Tube. A gas tube is a vacuum tube in which the pressure of the contained gas or vapor is such as to substantially affect the electrical characteristics of the tube.

Grid. A grid is an electrode having one or more openings for the passage of electrons or ions.

Grid Bias. Grid bias is the direct component of grid voltage.

Harmonic. A harmonic is a component of a periodic phenomenon having a frequency that is an integral multiple of the fundamental frequency. For example, a component the frequency

of which is *twice* the fundamental frequency is called the *second* harmonic.

Heater. A heater is an electrical heating element for supplying heat to an indirectly heated cathode.

Heptode. A heptode is a seven-electrode vacuum tube containing an anode, a cathode, a control electrode, and four additional electrodes ordinarily in the nature of grids.

Hexode. A hexode is a six-electrode vacuum tube containing an anode, a cathode, a control electrode, and three additional electrodes ordinarily in the nature of grids.

High-vacuum Tube. A high-vacuum tube is a vacuum tube evacuated to such a degree that its electrical characteristics are essentially unaffected by gaseous ionization.

Indirectly Heated Cathode. An indirectly heated cathode is a cathode of a thermionic tube to which heat is supplied by an independent heater element.

Input Circuit. The input circuit of a vacuum tube is that circuit associated with the control electrode.

Kilocycle. A kilocycle, when used as a unit of frequency, is 1000 cycles/sec.

Megacycle. A megacycle, when used as a unit of frequency, is 1,000,000 cycles/sec.

Mercury-vapor Tube. A mercury-vapor tube is a gas tube in which the active contained gas is mercury vapor. The term is most commonly applied to a diode, a mercury-vapor triode being known as a *thyratron*.

Modulated Wave. A modulated wave is a wave of which either the amplitude, frequency, or phase is varied in accordance with a signal.

Mu Factor. Mu factor (or μ factor) is the ratio of the change in one electrode voltage to the change in another electrode voltage, under the conditions that a specified current remains unchanged and that all other electrode voltages are maintained constant. It is a measure of the relative effect of the voltages on two electrodes upon the current in the circuit of any specified electrode. As most precisely used, the term refers to infinitesimal changes as indicated by the defining equation.

$$\mu_{jkl} = -\left(\frac{\partial e_j}{\partial e_k}\right), \; i_l \text{ constant}$$

Amplification factor is a special case of μ factor.

Multielectrode Tube. A multielectrode tube is a vacuum tube containing more than three electrodes associated with a single electron stream.

Mutual Conductance. See Control-grid-plate Transconductance.

Octode. An octode is an eight-electrode vacuum tube containing an anode, a cathode, a control electrode, and five additional electrodes ordinarily in the nature of grids.

Output Circuit. The output circuit of a vacuum tube is the circuit associated with the electrode from which useful power is drawn, generally the plate.

Pentode. A pentode is a five-electrode vacuum tube containing an anode, a cathode, a control electrode, and two additional electrodes ordinarily in the nature of grids.

Phototube. A phototube is a vacuum tube in which one of the electrodes is irradiated for the purpose of causing electron emission.

Plate. Plate is a common name for the principal anode in a vacuum tube.

Plate Resistance. Plate resistance is the quotient of the alternating plate voltage by the in-phase component of the alternating plate current, all other electrode voltages being maintained constant. As most precisely used, the term refers to infinitesimal amplitudes as indicated by the defining Eq. (3-24b). (See also pages 47 and 61.)

Radio Frequency. A radio frequency is a frequency at which radiation of electromagnetic energy, for communication purposes, is possible.

Rectification Factor. Rectification factor is the quotient of the change in average current of an electrode by the change in amplitude of the alternating sinusoidal voltage applied to the same electrode, the direct voltages of this and other electrodes being maintained constant. As most precisely used the term refers to infinitesimal changes. (See page 569.)

Screen Grid. A screen grid is a grid placed between a control grid and an anode and maintained at a fixed positive potential, for the purpose of reducing the electrostatic influence of the anode in the space between the screen grid and the cathode and of drawing electrons away from the cathode.

Side Band. A side band is a band of frequencies on either side of the carrier frequency, produced by the process of modulation. A side band may consist of a single frequency, in which case it is called a *side frequency*.

Signal. A signal is the form or variation with time of a wave whereby the information, message, or effect is conveyed in communication.

Space-charge Grid. A space-charge grid is a grid that is placed adjacent to the cathode and positively biased so as to reduce the limiting effect of space charge on the current through the tube.

Stage of Amplification. A stage of amplification consists of an amplifier tube together with its input and output circuits. Several stages may be connected in cascade to constitute a complete amplifier.

Suppressor Grid. A suppressor grid is a grid (usually connected electrically to the cathode) interposed between two electrodes (usually the screen grid and plate) both positive with respect to the cathode, in order to prevent the passing of secondary electrons from one to the other. It also serves to reduce the electrostatic influence of the anode in the space between the suppressor grid and the cathode.

Tank Circuit. A tank circuit is an oscillatory circuit so coupled to the output of a vacuum-tube amplifier or oscillator as to control the wave shape of the plate voltage by virtue of its energy storage.

Tetrode. A tetrode is a four-electrode vacuum tube containing an anode, a cathode, a control electrode, and an additional electrode ordinarily in the nature of a grid.

Thermionic Tube. A thermionic tube is a vacuum tube in which one of the electrodes is heated for the purpose of causing electron or ion emission from that electrode.

Transconductance. Transconductance from one electrode to another is the quotient of the in-phase component of the alternating current of the second electrode by the alternating voltage of the first electrode, all other electrode voltages being maintained constant. As most precisely used, the term refers to infinitesimal amplitudes. (See pages 61 and 67.)

Transrectification Factor. Transrectification factor is the quotient of the change in average current of an electrode by the change in the amplitude of the alternating sinusoidal voltage

applied to another electrode, the direct voltages of this and other electrodes being maintained constant. As most precisely used, the term refers to infinitesimal changes. (See page 569.)

Triode. A triode is a three-electrode vacuum tube containing an anode, a cathode, and a control electrode.

Vacuum Tube. A vacuum tube is a device consisting of an evacuated enclosure containing a number of electrodes, between two or more of which conduction of electricity through the vacuum or contained gas may take place.

Wave. A wave is (1) a propagated disturbance, usually periodic, as an electric wave or sound wave, (2) a single cycle of such a disturbance, (3) a periodic variation represented by a graph.

Wave Length. A wave length is the distance traveled in one period or cycle by a periodic disturbance. It is the distance between corresponding phases of two consecutive waves of a wave train. Wave length is the quotient of velocity by frequency.

SYMBOLS

The letter symbols used in this text are those recommended by the Standards Committee of IRE in their 1938 report, with a few minor exceptions and additions. (1) Instantaneous values of voltage and current are represented by small letters in conformance with general practice in all electrical engineering work. (2) Direct currents and voltages as well as effective and crest values of alternating currents and voltages are represented by capital letters. Crest (or maximum) values are designated by a subscript m. (3) Where it is necessary to distinguish vector or complex quantities, boldface type is used. (4) Subscripts are so chosen as to be as nearly indicative of the portion of the circuit to which they apply as possible. (5) Where any given quantity is to be restricted in use to a single frequency, it is shown with a subscript enclosed in parentheses to indicate the frequency, as $R_{1(p)}$, which indicates that the value of R_1 referred to is that obtained at a frequency corresponding to an angular velocity of p. (6) Occasions may arise where the average and quiescent (no alternating excitation) values may be different. In such cases the quiescent value may be distinguished by adding the subscript 0, as I_{b0}. (7) The power supplies for the various electrodes are indicated by the same symbols as are used for the average values except that the subscript is doubled, as E_{bb} for the plate power supply.

The information for Tables A-1 and A-2 is taken from the 1938 IRE Standards Report with some additional symbols that the author has found useful.

Figure A-1 illustrates the use of the more important of the volt-

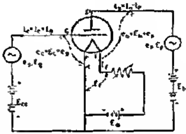


Fig. A-1.—Illustrating the application of the nomenclature to the various circuits of a triode tube.

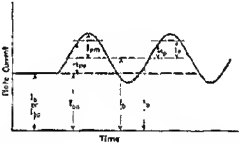


Fig. A-2.—Illustrating the application of the nomenclature to the plate current of a vacuum tube.

age and current symbols. The reason for using a minus sign between I_b and i_p is given on page 63. The minus sign is not used for the grid current or for the grid or plate voltages.

Figure A-2 further illustrates the use of the various symbols as applied to the plate current. Symbols for the grid current and the plate and grid voltages are treated similarly.

TABLE A-1

Quantity	Symbol	Subscript	Examples
Direct components	E I	b plate c (or $c1$) control grid $c2$ screen grid	E_b I_b E_{c1} E_{c2}
Quiescent values (if application of an alternating emf changes the direct component, these symbols are used to denote values before application of alternating emf.)	E I	$b0$ plate $c0$ grid	E_{b0} I_{b0}
Direct supply voltages (these symbols are used only when the supply voltage differs from the voltage applied to the tube)	E	bb plate cc grid	E_{bb} E_{cc}
Rms values of alternating components	E I	p plate g (or $g1$) control grid $g2$ screen grid	E_p I_p E_g I_{g1}
Crest values of alternating components	E_m I_m	p plate g (or $g1$) control grid $g2$ screen grid	E_{pm} I_{pm} E_{gm} I_{gm}
Instantaneous values of alternating components	e i	p plate g (or $g1$) control grid $g2$ screen grid	e_p i_p e_g i_g
Instantaneous total values (sum or difference of instantaneous direct and alternating components)	e i	b plate c (or $c1$) control grid $c2$ screen grid	e_b i_b e_{c1} e_{c2}
Instantaneous components above and below quiescent values (see footnote on page 488 for further explanation)	e i	$p0$ plate $c0$ grid	e_{p0} i_{p0}

TABLE A-2

Quantity	Symbol
Filament, or heater, terminal voltage	E_f
Filament, or heater, current	I_f
Total electron emission	I_s
Plate resistance (a-c)	$r_p = \frac{\partial e_b}{\partial i_b}$
Grid-plate transconductance (mutual conductance)	$g_m = g_{pe} = \frac{\partial i_b}{\partial e_c}$
Amplification factor	$\mu = -\frac{\partial e_b}{\partial e_c}$
Plate resistance (d-c)	R_b

TABLE A-3

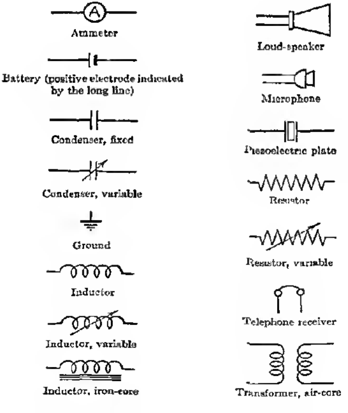


TABLE A-3.—*Continued*

Transformer, iron-core

Wires, joined

Voltmeter

Wires, crossed, not joined

Vacuum-tube Symbols

High-vacuum-tube envelope

Indirectly heated cathode

Gas-tube envelope
(dot placed where convenient)

Pool cathode

Directly heated cathode
(symbol often used in a general sense
to indicate a cathode of any type)

Photoelectric cathode

Grid

Cold cathode

Plate

TABLE A-4. ABBREVIATIONS

ampere	amp
amplitude-modulated	a-m
audio-frequency (adj.)	a-f
continuous waves	cw
decibel	db
electromotive force	emf
frequency-modulated	f-m
henry	h
high-frequency (adj.)	h-f
intermediate-frequency (adj.)	i-f
kilocycle (per second)	kc
low-frequency (adj.)	l-f
megacycle (per second)	Mc
megohm	MΩ
microfarad	μf
micromicrofarad	$\mu\mu$f
microhenry	μh
millihenry	mh
microampere	μa
milliampere	ma
radio-frequency (adj.)	r-f
root-mean-square (adj.)	rms
tuned radio-frequency (adj.)	t r f

Abbreviations for Metric Prefixes

centi	c
deci	d
deka	dk
hecto	h
kilo	k
mega	M
micro	μ
milli	m

APPENDIX B

FOURIER ANALYSIS OF A REPEATING FUNCTION

It may be shown that any periodic function may be represented by a series of sine waves, or

$$y = f(x) \tag{B-1}$$

may be written

$$y = a_0 + c_1 \sin(\omega t + \phi_1) + c_2 \sin(2\omega t + \phi_2) + c_3 \sin(3\omega t + \phi_3) + \cdots \tag{B-2}$$

provided that $y = f(x)$ is a repeating function of period $2\pi/\omega$. Equation (B-2) may evidently be written as

$$y = a_0 + a_1 \cos \omega t + a_2 \cos 2\omega t + a_3 \cos 3\omega t + \cdots + b_1 \sin \omega t + b_2 \sin 2\omega t + b_3 \sin 3\omega t + \cdots \tag{B-3}$$

where

$$c_1 = \sqrt{a_1^2 + b_1^2}, \qquad c_2 = \sqrt{a_2^2 + b_2^2}, \qquad \cdots$$

and

$$\phi_1 = \tan^{-1} \frac{a_1}{b_1}, \qquad \phi_2 = \tan^{-1} \frac{a_2}{b_2}, \qquad \cdots$$

The constants a_0, a_1, b_1, a_2, b_2, etc., may be evaluated by the following procedure:

To determine a_0, find the average of y over a time interval of $2\pi/\omega$, or

$$\frac{1}{2\pi} \int_0^{2\pi} y \, d\omega t = \frac{1}{2\pi} \int_0^{2\pi} (a_0 + a_1 \cos \omega t + a_2 \cos 2\omega t + \cdots + b_1 \sin \omega t + b_2 \sin 2\omega t + \cdots) \, d\omega t$$

Evidently the integration of the right-hand side will give a_0, so that

$$\frac{1}{2\pi} \int_0^{2\pi} y \, d\omega t = a_0 = \text{average of } y \text{ over one cycle} \tag{B-4}$$

a_1 may be determined by first multiplying both sides of Eq. (B-3) by $\cos \omega t$ and again integrating over a cycle,

$$\frac{1}{2\pi}\int_0^{2\pi} y \cos \omega t \, d\omega t$$
$$= \frac{1}{2\pi}\int_0^{2\pi} (a_0 \cos \omega t + a_1 \cos^2 \omega t + a_2 \cos \omega t \cos 2\omega t + \cdots + b_1 \cos \omega t \sin \omega t + b_2 \cos \omega t \sin 2\omega t + \cdots) \, d\omega t$$

Integration of the right-hand side gives

$$\frac{1}{2\pi}\int_0^{2\pi} y \cos \omega t \, d\omega t = \frac{a_1}{2}$$
$$= \text{average of } (y \cos \omega t) \text{ over one cycle} \qquad \text{(B-5)}$$

b_1 may be determined by multiplying both sides of Eq. (B-3) by $\sin \omega t$ and integrating, giving

$$\frac{1}{2\pi}\int_0^{2\pi} y \sin \omega t \, d\omega t = \frac{b_1}{2}$$
$$= \text{average of } (y \sin \omega t) \text{ over one cycle} \qquad \text{(B-6)}$$

a_2 is determined by first multiplying by $\cos 2\omega t$, and b_2 by multiplying by $\sin 2\omega t$. The process may be continued to include any number of harmonics or any particular harmonic desired.

In the analysis of a wave by this method the ordinates y are measured at frequent intervals and tabulated. The number of intervals taken per cycle depend upon the accuracy of the results desired, although there is no point in taking them closer than the accuracy of the original curve warrants.

Table B-1 illustrates the application of this method to the analysis of the plate-current curve of Fig. 9-46. Column 1 gives the intervals taken along the abscissa, in this case every 10 deg. Column 2 gives the measured values of the ordinate [y in Eq. (B-2)] Column 3 gives the product of column 2 and $\cos \omega t$, ($y \cos \omega t$), etc. The process of integration consists in totaling these columns as shown and then dividing by the number of intervals, (36), to get the average. a_0, a_1, b_1, etc., are, therefore, given by Eqs. (B-4), (B-5), (B-6), etc., within the accuracy of the original curve and the approximation involved in using finite intervals to evaluate the integrals (sometimes called step-by-step integration). In this

TABLE B-1. FOURIER ANALYSIS OF CURVE OF FIG. 9-46

(1) ωt	(2) i_b	(3) $i_b \cos \omega t$	(4) $i_b \sin \omega t$	(5) $i_b \cos 2\omega t$	(6) $i_b \sin 2\omega t$	(7) $i_b \cos 3\omega t$	(8) $i_b \sin 3\omega t$
0	64.5	64.5	0	64.5	0	64.5	0
10	64	63.0	11.1	60.2	21.9	55.5	32.0
20	63	59.2	21.5	48.3	40.5	31.5	54.6
30	60	52.0	30.0	30.0	52	0	60.0
40	56.5	43.3	36.3	9.9	55.6	−28.2	49.0
50	52	33.4	39.8	−9.0	51.2	−45.0	26.0
60	47	23.4	40.7	−23.4	40.7	−46.6	0
70	41.5	14.2	39.0	−31.8	26.7	−35.8	−20.7
80	34	5.9	33.5	−32.0	11.6	−17.0	−30.4
90	29	0	29.0	−29.0	0	0	−29.0
100	24	−4.1	23.6	−22.5	−8.2	11.9	−20.8
110	10.5	−6.7	16.9	−14.9	−12.5	16.9	−9.8
120	15.5	−7.7	13.5	−7.7	−13.5	15.5	0
130	11	−7.1	8.4	−1.9	−10.8	9.5	5.5
140	8.5	−6.5	5.5	1.4	−8.4	4.3	7.8
150	6	−5.2	3.0	3.0	−5.2	0	6.0
160	4.5	−4.1	1.5	3.4	−2.9	−2.2	3.9
170	3.5	−3.4	0.6	3.3	−1.2	−3.1	1.7
180	3	−3.0	0	3.0	0	−3.0	0
190	3.5	−3.4	−0.6	3.3	1.2	−3.1	−1.7
200	4.5	−4.1	−1.5	3.4	2.9	−2.2	−3.0
210	6	−5.2	−3.0	3.0	5.2	0	−6.0
220	8.5	−6.5	−5.5	1.4	8.4	4.3	−7.8
230	11	−7.1	−8.4	−1.9	10.8	9.5	−5.5
240	15.5	−7.7	−13.5	−7.7	13.5	15.5	0
250	19.5	−6.7	−16.9	−14.9	12.5	16.9	9.8
260	24	−4.1	−23.6	−22.5	8.2	11.9	20.8
270	29	0	−29.0	−29.0	0	0	29.0
280	34	5.9	−33.5	−32.0	−11.6	−17.0	30.4
290	41.5	14.2	−39.0	−31.8	−26.7	−35.8	20.7
300	47	23.4	−40.7	−23.4	−40.7	−46.6	0
310	52	33.4	−39.8	−9.0	−51.2	−45.0	−26.0
320	56.5	43.3	−36.3	9.9	−55.6	−28.2	−49.0
330	60	52.0	−30.0	30.0	−52	0	−60.0
340	63	59.2	−21.5	48.3	−40.5	31.5	−54.6
350	64	63.0	−11.1	60.2	−21.9	55.5	−32.0
Totals	1146.5	560.7	0	42.1	0	−4.1	0

$$a_0 = \frac{1146.5}{36} = 31.8, \qquad a_1 = \frac{560.7}{18} = 31.1,^* \qquad a_2 = \frac{42.1}{18} = 2.34$$

$$a_3 = \frac{-4.1}{18} = -0.23, \quad b_1 = 0, \quad b_2 = 0, \quad b_3 = 0$$

$$a_0 = I_l, \quad \sqrt{a_1^2 + b_1^2} = I_{1m}, \quad \sqrt{a_2^2 + b_2^2} = I_{2m}, \quad \sqrt{a_3^2 + b_3^2} = I_{3m}$$

* The totals of columns (3), (4), etc., are divided by 18 rather than 36 because the averages of these columns are $a_1/2$, $b_1/2$, ... Thus it is necessary to divide by 36 and then multiply by 2 to get a_1, b_1, . . .

particular example the curve is symmetrical about the Y-axis, and all the b terms are zero.

It should be noted that this method may be used to evaluate any given harmonic without first evaluating those of lower order. For example, if the third harmonic only was desired, columns 1, 2, 7, and 8 alone would be necessary.

APPENDIX C

APPLICATION OF THE FOURIER ANALYSIS TO AN ANALYTICAL SOLUTION OF REPEATING FUNCTIONS, THE EQUATIONS OF WHICH ARE KNOWN OVER SHORT INTERVALS

Certain types of repeating functions for which the equations are known over short, well-defined intervals may be analyzed by the use of actual integration processes rather than by the step-by-step method of Appendix B. An example is given here.

Example. The plate current of a class B amplifier closely approximates a series of half-sine waves, Fig. C-1 (also see pages 354 and 395). The equation of this curve may be written for the region between points a and b, as $i = I_m \sin \omega t$, and between b and a' as $i = 0$.

FIG. C-1.—Theoretical plate-current curve in a class B amplifier. Each pulse is an exact half-sine wave.

The constant a_0 of Eq. (B-3), Appendix B, may be evaluated by applying Eq. (B-4) analytically. Summation of all the ordinates, as in column 2, Table B-1, is performed by integrating $i = I_m \sin \omega t$ from a to b, or from 0 to π, since i is zero between π and 2π. a_0 is then given by dividing this summation by the total number of ordinates, or 2π. Therefore,

$$a_0 = \frac{1}{2\pi} \int_0^\pi I_m \sin \omega t \, d\omega t = 0.318 I_m$$

Similarly, from Eq. (B-5),

$$a_1 = 2 \times \frac{1}{2\pi} \int_0^\pi I_m \sin \omega t \cos \omega t \, d\omega t = 0$$

and

$$a_2 = 2 \times \frac{1}{2\pi} \int_0^\pi I_m \sin \omega t \cos 2\omega t \, d\omega t = -0.212 I_m$$

$$b_1 = 2 \times \frac{1}{2\pi} \int_0^\pi I_m \sin \omega t \sin \omega t \, d\omega t = 0.5 I_m$$

$$b_2 = 2 \times \frac{1}{2\pi} \int_0^\pi I_m \sin \omega t \sin 2\omega t \, d\omega t = 0$$

Higher frequency components may be determined in the same manner.

APPENDIX D

METHODS OF EVALUATING THE NUMERICAL VALUES IN TABLE 7-1

All the numerical values in Table 7-1 (page 195) may be evaluated by reasonably simple, straightforward methods. They are outlined herewith for a three-phase, half-wave rectifier, the circuit of which was shown in Fig 7-8 (page 169).

A.1. Ratio of Transformer Secondary Voltage (rms) to Direct Output Voltage. Since all d-c drop through the rectifier circuit is neglected in computing Table 7-1, the direct voltage will evidently be the average of the output voltage that the rectifier

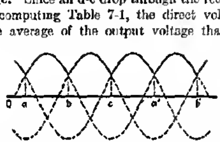


FIG D-1 —Output wave of a three-phase, half-wave rectifier The heavy line is the output voltage, the dotted lines showing the impressed voltages of each of the three phases (tube and transformer drop neglected)

applies to the filter, as shown in Fig. 7-9. The ratio of this average to the crest value of the alternating voltage may be obtained by the methods of Appendixes B and C.

Consider the output voltage of the rectifier, as shown by the solid line of Fig. D-1. The equation of the curve between a and b is $e = E_m \sin qt$ (q is used here rather than ω to avoid later confusion). Between b and c and between c and a' the curve is the same but displaced an additional 120 deg. Evidently, too, the period ab is one full cycle of the rectified output so that the integrations indicated by Eqs. (B-4) to (B-6) must be extended from a to b or over any other similar section That is, the fundamental frequency of the output wave is three times that of the supply, or $\omega = 3q$ [where ω has the significance assigned to it by Eq (B-2)]

Equation (B-4) therefore becomes

$$a_0 = \frac{1}{2\pi}\int_a^b E_m \sin qt \, d\omega t \tag{D-1}$$

where $a = \pi/2$, $b = 5\pi/2$.

Integration of Eq. (D-1) gives $a_0 = 0.828E_m$, or

$$E_s = 0.855E_0$$

where $E_s = 0.707E_m$ and $E_0 = a_0$. 0.855 is the figure given in Table 7-1 for the ratio of the rms value of the secondary voltage to the direct output voltage.

A.2. Ratio of Inverse Peak Voltage to Direct Output Voltage. The inverse peak voltage is the maximum voltage applied across the tube in a negative, or inverse, direction. It may be determined from the curve of output voltage (Fig. 7-9) reproduced in Fig. D-2*a*. Referring to the discussion that accompanied Fig. 7-9, the anode potential of, say, tube 1 is given by the sine wave (partly dotted) marked 1, and the cathode potential is given by the solid line envelope. The instantaneous inverse voltage is then given by the difference between these two curves as illustrated by the shaded area in curve (*a*) Fig. D-2 and by the curve (*b*) of the same figure. The inverse peak voltage is therefore found by determining the maximum difference between the upper envelope and curve 1. This may be done by writing the equation for each curve over the interval t_2t_3 (or t_2t_1'), subtracting the two equations, differentiating with respect to qt, and equating the derivative to zero. The equation of curve 1 may be written

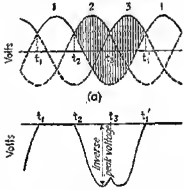


FIG. D-2.—Inverse voltage curves. The curves in (*a*) are similar to those of Fig. 7-9 (p. 169) and have the same significance. Curve (*b*) shows the actual voltage across tube 1 at each instant, being obtained from the curves in (*a*). The conducting period is from t_1 to t_2 and the nonconducting period from t_2 to t_1'.

$$e_1 = E_m \sin qt$$

and that of the envelope for the interval t_2t_3 is

$$e = E_m \sin\,(qt - 120°)$$

The instantaneous inverse voltage during this interval is then

$$\text{Inverse voltage} = E_m \sin qt - E_m \sin\,(qt - 120°) \qquad \text{(D-2)}$$

Taking the derivative of this voltage with respect to t and equating to zero yields

$$\frac{\sqrt{3}}{2} E_m \sin qt - \frac{3}{2} E_m \cos qt = 0$$

or $qt = 240°$. This value of qt gives the point on the curve of Fig. D-2 at which the inverse voltage is a maximum and if inserted in Eq. (D-2) will give the inverse peak voltage.

$$E_{\text{inv}} = E_m \sin 240° - E_m \sin 120° = -\sqrt{3}\, E_m$$

the minus sign merely indicating that the plate voltage is negative. From the preceding section we find that $E_m = E_0/0.828$, so we may write

$$E_{\text{inv}} = \frac{\sqrt{3}\, E_0}{0.828} = 2.09E_0$$

which is the value given in Table 7-1. (The minus sign is omitted, since it is obvious that the inverse voltage must be negative.)

A.3. Ratio of First Three Alternating Components of Rectifier Output (rms), to the Direct Output Voltage. The alternating components of the output voltage may be determined by continuing the Fourier analysis method used in part A.1

Applying this method to Eq. (B-5) gives the fundamental component of the ripple voltage in the output

$$\frac{a_1}{2} = \frac{1}{2\pi} \int_{\pi/2}^{5\pi/2} E_m \sin qt \cos \omega t \, d\omega t$$

where $\omega = 3q$ as before. Integration gives $a_1 = 0$.

Equation (A-6) gives

$$\frac{b_1}{2} = \frac{1}{2\pi} \int_{\pi/2}^{5\pi/2} E_m \sin qt \sin \omega t \, d\omega t$$

which, when integrated, gives

$$b_1 = -0.207E_m = -0.207\,\frac{E_0}{0.828} = -0.25E_0$$

Since $a_1 = 0$, c_1 of Eq. (B-2) (page 613) is numerically equal to b_1. The crest value of the fundamental component is equal to c_1; therefore, the rms voltage is given by

$$E_1 = 0.707(0.25E_0) = 0.177E_0$$

where E_1 is the rms value of the fundamental component, and E_0 is the d-c, or average, value.

The method may be applied to higher order harmonics in a similar manner.

B. Ripple Frequency. The ripple frequency may be readily determined by inspection of the output voltage curve, Fig. 7-9.

C.1. Ratio of Peak Anode Current to Direct Output Current. In the three-phase, half-wave rectifier only one tube conducts at a time. Therefore, the peak current (under the assumption of infinite filter inductance) must be equal to the direct current. Where two tubes conduct simultaneously *in parallel* (not in series obviously) as in the double-Y circuit, the peak current is half the direct current.

C.2. Ratio of Average Anode Current to Direct Output Current. The conduction period for the three-phase, half-wave rectifier is one-third of a cycle. Therefore, the average tube current is one-third of the peak current and therefore one-third of the direct output current.

D.1. Primary Utilization Factor. The wave shape of the current flowing through the transformers of a rectifier is obviously far from sinusoidal; therefore, the heat losses will be greater than under sinusoidal operation. Transformers, like other electrical equipment, are rated in terms of the output that they will deliver with a temperature rise not to exceed a predetermined safe value, *under the assumption that they will be called upon to carry sine-wave currents.* Ratings given on such a basis must therefore be altered for transformers used in rectifier circuits, being multiplied by a term known as the *transformer utilization factor.* This factor may be defined as the ratio of the output of the transformer when used in a given rectifier circuit to the output that it would be capable of giving with sinusoidal currents flowing *and the same internal loss.* Evaluation of this factor therefore requires that the heat losses be determined in the transformer when in rectifier service.

The impressed emf across a transformer is ordinarily sinusoidal regardless of the type of load. Thus the flux is nearly sinusoidal,

and the iron losses may be considered as independent of the wave shape of the load current. The copper losses, on the other hand, are directly dependent upon the current flowing through the windings and are therefore markedly affected by the wave shape. Under the assumptions of Table 7-1, *viz.*, infinite inductance in the filter, the secondary current for the 3-ϕ, half-wave rectifier will be a rectangular pulse lasting for one-third of a cycle and having an amplitude equal to the direct load current (curve *a*, Fig. D-3). Since the transformer is considered ideal, the primary current (curve *b*, Fig. D-3) must be of the same shape except that no direct component can be present, since there is no source of direct emf in the primary circuit.

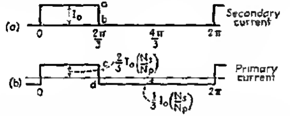


FIG. D-3.—Primary and secondary currents in one transformer of a three-phase half-wave rectifier, assuming an infinite-inductance choke and an ideal transformer.

Proof that the curve *b* of Fig. D-3 correctly represents the alternating current may best be handled in two steps. First, it is evident that the secondary current consists of a direct component, equal to the average current taken over one cycle, with a superimposed alternating component. Since only the alternating component can induce an emf into the primary, the average primary current must be zero, as shown in curve *b*; *i.e.*, it contains no direct component. The second step is to show that the primary current in an ideal transformer is rectangular in shape, which may be seen by first noting that the main transformer flux is sinusoidal (see preceding paragraph). Thus under no-load conditions the only current flowing will be the magnetizing current in the primary winding which, in an ideal transformer, is negligibly small. When a load is applied, any variations in current in the secondary winding will tend to vary the wave shape of the flux and so alter the wave shape of the induced emf in the primary winding. Since the

primary induced emf must always equal the impressed emf (in an ideal transformer), it is evident that the ampere turns due to this *changing* secondary current must be offset by an equal number of ampere turns of opposite polarity supplied by an additional current flowing in the primary in order that the net excitation may remain that of the exciting current only. Thus the primary current must have the same wave shape as the secondary but without a direct component. Since the only points in the cycle at

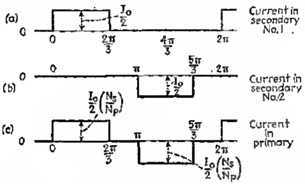


FIG. D-4.—Primary and secondary currents in one transformer of a three-phase full-wave, double-Y rectifier, assuming an infinite-inductance choke and an ideal transformer.

which the secondary current changes are at $qt = 0$ and $qt = 2\pi/3$, we may write

$$N_p(\Delta I_1) = N_s(\Delta I_2)$$

where (ΔI_1) is given by *cd* and (ΔI_2) by *ab*, Fig. D-3, and N_p and N_s are the number of turns in the primary and secondary windings respectively. The current change *ab* is equal to I_0; therefore, *cd* must be $I_0(N_s/N_p)$. In order that the average current may be zero, two-thirds of this change must be above the zero line and one-third below as shown in curve *b*, Fig. D-3.[1]

[1] In full-wave circuits, of course, no d-c magnetization whatsoever exists in the core. As an example, the wave shapes of the primary and secondary currents for the $3\text{-}\phi$, double-Y circuit of Fig. 7-13, with infinite filter inductance, are shown in Fig. D-4. Curves *a* and *b* represent the currents in each of the two secondaries of one transformer. The magnetization of each is opposite in direction to that of the other and thus produces zero average (or d-c) flux. To indicate this relation, the current in curve *b* is shown as a

From the definition of utilization factor we may now write

$$\text{Utilization factor} = \frac{E_0 I_0/3}{E_p I_p} \tag{D-3}$$

where $E_0I_0/3$ is the output demanded of the transformer[1] under the assumptions applying to Table 7-1 (no loss in the rectifier circuit) and E_pI_p is the sinusoidal rating of the transformer. E_p is, of course, the actual primary voltage of the transformer, so I_p is evidently the sinusoidal current that will produce the same copper losses as are produced by the actual current flowing in rectifier service.

If the resistance of the transformer primary winding is R_p, the copper losses produced by I_p are

$$P_{pn} = I_p^2 R_p \tag{D-4}$$

where P_{pn} = copper loss in primary under normal sinusoidal conditions. The loss under rectifier operation may be found by determining the energy consumed during each of the two intervals 0 to $2\pi/3$ and $2\pi/3$ to 2π, adding them and dividing the total by the time of one cycle to secure the average power, thus:

$$P_{pr} = \frac{1}{T}\left[\left(\frac{2}{3} I_0 \frac{N_s}{N_p}\right)^2 R_p \frac{T}{3} + \left(\frac{I_0}{3}\frac{N_s}{N_p}\right)^2 R_p \frac{2T}{3}\right] = \frac{2}{9} I_0^2 R_p \left(\frac{N_s}{N_p}\right)^2 \tag{D-5}$$

where P_{pr} = copper loss in primary under rectifier conditions
T = time of one cycle
R_p = resistance of primary winding

By the definition of the utilization factor $P_{pn} = P_{pr}$, or

$$I_p^2 R_p = \frac{2}{9} I_0^2 R_p \left(\frac{N_s}{N_p}\right)^2 \tag{D-6}$$

which, when solved for I_p, yields

$$I_p = 0.472 I_0 \frac{N_s}{N_p} \tag{D-7}$$

---

negative quantity, although it flows in the same direction as that of c insofar as the external d-c circuit is concerned. The primary current may now readily be seen to have the shape of curve c.

[1] Division by 3 is necessary because there are three transformers supplying the total output, E_0I_0.

Since $E_s = 0.855E_0$ from part A.1 of this appendix, and since $E_p = \frac{N_p}{N_s} E_s$ (where E_s is the voltage of the transformer secondary), we may rewrite Eq. (D-3)

$$\text{Utilization factor} = \frac{E_0 I_0}{3E_p I_p} = \frac{E_0 I_0}{3(0.855E_0)(0.472I_0)} = 0.827$$

D.2. Secondary Utilization Factor. The method of determining this factor is exactly the same as for the primary. In this case the secondary copper loss under rectifier conditions will be

$$P_{sr} = \frac{1}{T}\left(I_0^2 R_s \frac{T}{3}\right) = \frac{I_0^2 R_s}{3} \tag{D-8}$$

and under sinusoidal conditions

$$P_{ss} = I_s^2 R_s \tag{D-9}$$

Setting $P_{ss} = P_{sr}$, I_s is found to be

$$I_s = 0.578 I_0 \tag{D-10}$$

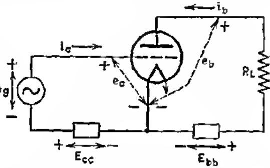
Therefore,

$$\text{Utilization factor} = \frac{E_0 I_0}{3E_s I_s} = \frac{E_0 I_0}{3(0.855E_0)(0.578I_0)} = 0.675$$

APPENDIX E

RELATIVE POLARITIES OF CURRENTS AND VOLTAGES IN A VACUUM-TUBE AMPLIFIER

The circuit diagram of a simple vacuum-tube amplifier is shown in Fig. E-1, similar to Fig. 3-22, page 58. A plus sign and a minus sign are used to indicate that the plate voltage, e_b, and the grid voltage, e_c, are considered to be positive when acting from cathode



E-1. Circuit of a simple vacuum-tube amplifier

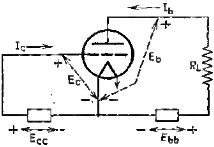
toward plate and grid, respectively, and an arrow is used as a similar indication for the plate current, i_b, and the grid current, i_c. Under these assumptions the drop, i_bR_L, across the load resistance must be considered as positive when acting from top to bottom of the resistor, as indicated by the (+) and (−) signs. The equation for the plate circuit is then

$$E_{bb} - i_bR_L - e_b = 0 \tag{E-1}$$

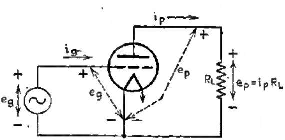
If the alternating voltage applied to the grid circuit is assumed to be zero ($e_g = 0$), only direct components of current and voltage will appear in the plate circuit and we may replace i_b with I_b and e_b with E_b, as in Fig. E-2. The equation of the plate circuit under these conditions is then

$$E_{bb} - I_bR_L - E_b = 0 \tag{E-2}$$

Let us now suppose that an alternating voltage is again applied to the grid but let us neglect the direct components and show a circuit containing the alternating components only, as in Fig. E-3. Here the plate voltage, e_b, must be replaced by the alternating component only, e_p. It seems logical to assume the same polarity, so the voltage e_p is shown with (+) and (−) signs indicating the same positive direction as for e_b in Fig. E-1. It is



E-2. Same as Fig. E-1, except for direct components only.



E-3. Same as Fig. E-1, except for alternating components only.

further evident that the potential across the load resistor is the same as that between plate and cathode and is, therefore, e_p, as indicated on the diagram. If the alternating plate current, i_p, is now assumed to be positive when flowing into the plate, as was assumed for i_b in Fig. E-1, we can see that we would have to write, for the potential across R_L, $e_p = -i_pR_L$. While the use of the minus sign in front of i_pR_L is permissible, it seems awkward and may be avoided by considering that the alternating plate current is positive when flowing *away from* the plate. This means that e_b and i_b are given by

$$e_b = E_b + e_p \tag{E-3}$$

$$i_b = I_b - i_p \tag{E-4}$$

The power relations in the circuit may be determined by noting that any circuit element is assumed to be a source of power (delivers power *into* the circuit) if the positive direction of the current is *out of* the (+) terminal of the potential, while it is a sink of power (absorbs power *from* the circuit) if the positive direction of the current is *into* the (+) terminal of the potential. Thus in Figs. E-1 and E-2 the direct potential, E_{bb}, is treated as a source of power and the power delivered to each circuit is $E_{bb}i_b$ and $E_{bb}I_b$, respectively. Similarly both the plate of the tube and the load resistance R_L are treated as though they were sinks, absorbing power in the amount of e_bi_b and $i_b^2R_L$ respectively in Fig. E-1 and E_bI_b and $I_b^2R_L$ in Fig. E-2.

In Fig E-3 the positive direction of current is out of the (+) terminal of the potential between plate and cathode of the tube indicating that the plate is assumed to be a *source* of power, rather than a sink as in the two preceding figures Thus the plate delivers a power e_pi_p to the output circuit. The load resistor R_L is again considered as a sink absorbing a power $i_p^2R_L$ or e_pi_p. This is typical of an amplifier where the plate is the apparent source of the a-c power delivered to the external load R_L.

An entirely different situation exists in the grid circuit. The alternating voltage is not normally generated by passing the alternating grid current through a resistance (or other type of impedance) but is produced by some external source. In other words the alternating voltage appearing in the external grid circuit is that of a *source* of power, not that of a *sink* as in the plate circuit. The normal assumption to make concerning the alternating component of the grid current is then to consider the positive direction as being toward the grid, as in Fig. E-3. This shows that we are considering the external circuit to be a source of power since the positive direction of the current is *out of* the (+) terminal of the applied potential. We then consider the grid-to-cathode circuit of the tube to be a sink, since the positive direction of the grid current is *into* the (+) terminal of the grid-to-cathode potential.

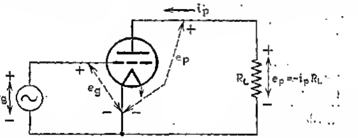
It is of interest to note that the positive direction of the direct grid potential, E_{cc}, is shown as being toward the grid, to correspond with the assumption for E_{bb}. Actually most tubes are operated

with a negative potential on the grid, and therefore, if a direct potential of, say, 20 volts is to be supplied by E_{cc}, we must write $E_{cc} = -20$ volts to indicate the correct absolute polarity. Many authors consider E_{cc} as positive when acting toward the *cathode* and so would write $E_{cc} = 20$ volts, but the notation presented here and used throughout this book seems to the author to be more consistent and less apt to result in errors or confusion.

It is possible to treat the plate circuit in the same manner as the grid circuit and consider the external voltage across the load impedance as that of a source of power and the plate-to-cathode circuit as a sink. This would be the result if the plate current arrow in Fig. E-3 were reversed, as in Fig. E-4 whence Eq. (E-4) would become

$$i_b = I_b + i_p \tag{E-5}$$

The positive direction of the current would then be *into* the (+) terminal of the plate-to-cathode potential and the plate would be a



E-4. Same as Fig. E-3 except that positive direction of i_p is assumed toward plate.

sink, absorbing an amount of power $e_p i_p$. When the relative phases of voltage and current are determined (as on pages 64–66) it will be found that e_p and i_p are out of phase by 180 deg when the plate current and plate voltage have the relative polarities indicated in Fig. E-4, and we must conclude that, while the plate of the tube is a sink, it is absorbing *negative* a-c power which is the equivalent of a *source* delivering *positive* power, exactly as was assumed in setting up the polarities of Fig. E-3.

Similarly the resistance R_L, if the plate current direction is as shown in Fig. E-4, is treated as a *source* of power, just as was the external circuit of the grid. Since $e_p = -i_p R_L$ under the direc-

tions assumed in Fig. E-4, the power *delivered* by the resistance to the rest of the circuit may be written $e_p i_p = (-i_p R_L) i_p = -i_p^2 R_L$. The negative sign indicates that the resistance is actually *absorbing* an amount of power $+i_p^2 R_L$ as, of course, it must.

The purpose of the discussion in the last two paragraphs is to show that the use of the minus sign in expanding i_b into $I_b - i_p$, as was done on page 63, is not an essential step but one that produces more normal power equations in an amplifier. If the vacuum tube is used in an application where an external source of a-c power is applied to the plate circuit so that the plate-to-cathode circuit becomes an actual sink, the use of the minus sign in the plate-current expression would be less desirable than that of the plus sign However such applications are rare, and even when they do occur there is no more difficulty in using the minus sign in the equation for i_b than there would be in using the plus sign in the usual amplifier applications, as illustrated in the preceding paragraphs

It should be pointed out that this problem is met in various ways by different authors. Some write Eq (E-3) with a minus sign and Eq. (E-4) with a plus sign. This produces positive power flow equally as well as do Eqs. (E-3) and (E-4) Others, especially more advanced writers publishing in technical periodicals, simply ignore the sign, since the direction of power flow is well known and they do not need to rely upon a plus or minus sign to show that a given power flow is that of a source or sink. Beginning students, on the other hand, are likely to become confused unless definite standards are set up and adhered to.

AUTHOR INDEX

A

Alexanderson, E. F. W., 247
Armstrong, E. H., 549
Atlee, Z. S., 158

B

Bahls, W. E., 128
Ballantine, S., 566
Barton, L. E., 348
Becker, S., 364
Bedford A. V., 154
Bedford, B. D., 247
Bennon, S., 572
Bereskin, A. B., 224, 294
Black, H. S., 361
Black, L. J., 311
Bode, H. W., 367
Breyer, J., Jr., 247
Brown, H. D., 245
Brown, R. H., 430
Brunetti, C., 479
Builder, G., 363
Burnap, R. S., 83
Buss, R. R., 521

C

Carter, P. S., 476
Chambers, D. E., 248
Child, C. D., 44
Clapp, J. K., 482
Cook, E. G., 313
Coroniti, S. C., 233
Craig, P. H., 247

D

Dailey, H. I., 32
Dawson, J. W., 247
Dickieson, A. C., 313
Dishal, M., 393
Doba, S., 313
Doherty, W. H., 432, 556
Dow, W. G., 93
Dreyer, J. F., 83
Durand, S. R., 239
Dushman, S., 14

E

Eastman, A. V., 250, 293, 533, 548
Edison, T. A., 5
Edwards, M. A., 247
Einstein, A., 19
Elder, F. R., 247
Epstein, D. W., 146
Esperson, G. A., 29
Everitt, W. L., 267, 385, 405

F

Farrington, J. F., 457
Ferns, J. H., 413
Fink, D. G., 153
Fisher, S. T., 436
Fitzpatrick, G. M., 117
Foos, C. B., 236
Freeman, R. L., 203, 457
Fremlin, J. H., 58
Fyler, G. W., 484

G

Garman, G. W., 248
Geiger, P. H., 210
George, E. E., 251
Gessford, R. K., 203
Glauber, J. J., 559, 600
Gleason, C. H., 204

Glover, A M , 18
Goldberg, H , 309
Gorham, J. E., 16, 29
Gray, T. S., 247
Green, C. B , 534
Grieg, D D., 559, 600
Griffith, R C , 247
Grondahl, L O , 210
Gustafson, W. G., 302

H

Hanna, C. R , 191
Hansell, C. W , 476
Harris, W. A., 534
Hergenrother, R. C., 203
Herold, E W., 58, 584
Herskind, C C., 236, 247
Hertz, H. R., 18
Hickok, W H , 155
Hill, W. R., Jr., 224, 226
Holland, R. N., 58
Horle, L. C F., 533
Horton, J W., 248
Huber, W A , 373
Hull, A. W., 99, 100
Hull, L. M., 432

I

Iams, H., 155, 157

J

Jackson, J. S , 485
Janes, R B., 155
Jones, M. C., 436

K

Kauzmann, A. P., 201
Kell, R D., 154
Keller, O., 239
Kilgour, C. E , 337
Kimball, C. N., 236
Klemperer, H , 210
Knowles, D. D., 208
Koehler, G., 348
Koller, L. R., 11, 93, 116
Koomans, N., 536
Korn, G. E., 153
Kozanowski, H. N., 405

L

Lamb, A., 141
Langmuir, I., 10, 109
Law, H., 157
Law, R. R., 153
Learned, V., 367
Lederer, E. A., 35
Lee, R , 191
Levine, A. M., 559, 600
Llewellyn, F. B , 459, 475, 490
Loftin, E H , 308
Lowry, E F , 208
Ludwig, L R , 112

M

McLean, T , 348
McLean, W. R , 235
McProud, C. G , 313
Maloff, I. G., 140, 153
Malter, M , 152, 153
Mason, W. P , 470, 472
Mathes, R. O., 313
Mautner, R. S , 203
Meacham, L. A., 475
Millikan, R., 20
Mittag, A H., 247
Morton, G A , 152, 155
Moskowitz, S., 559, 600
Mouromtseff, I. E., 31, 32, 405

N

Nelson, J R., 313
Nesslage, G F , 584
Nottingham, W. B., 323
Nyquist, H., 361, 365

O

Offner, F. F., 309
Overbeck, W. P., 239

P

Pan, Wen-Yuan, 367
Pickles, S. B., 182
Planck, M., 19
Pollack, R. S., 313
Preisman, A., 343
Prince, D. C., 106, 404, 459
Pullen, K. A., 439

R

Rajchman, R. A., 153
Rank, D. H., 153
Reich, H. J., 93, 375
Richardson, O. W., 13
Rider, J. F., 307
Roake, W. C., 405
Roberts, F., 572
Robinson, H. A., 521
Roder, H., 509, 537
Rose, A., 157
Rose, G. M., Jr., 89
Ryder, J. D., 93

S

Sabbuh, C. A., 243
Salzberg, B., 58
Sanford, F. E., 247
Schade, O. H., 89, 202, 203
Schlesinger, K., 375
Schmidt, A., Jr., 247
Schneider, E. D., 247
Schroeder, A. C., 439
Schultz, E. H., 226
Scott, E. D., 539, 548
Scott, H. H., 313
Scott, H. J., 311, 538
Shackleford, C. L., 204
Shapiro, D. L., 373
Shatford, P. A., 58
Shaw, A. E., 118
Slepian, J., 112, 210
Spitzer, E. E., 34
Stewart, H. H., 313
Stout, M. B., 185
Strong, C. E., 438
Sziklai, G. C., 439

T

Taskin, H. A. T., 204
Terman, F. E., 182, 363, 367, 405, 413, 476, 483, 521, 525, 572, 592
Thomas, C. H., 128
Thomas, H. P., 410
Thompson, B. J., 89, 343, 353
Thomson, J. J., 5
Tolmie, J. R., 53
Tompkins, F. N., 241
Trainor, M. A., 154
Travis, C., 595

V

van der Pol, B., 537
van der Wyck, C. T. F., 530
Van Dyke, K. S., 473
Van Scoyoc, J. N., 220
Vogdes, F. B., 106, 404, 459
von Ardenne, M., 338

W

Wagener, W. G., 354, 405, 432, 516 520
Waidelich, D. L., 204
Wamsley, D. H., 35
Weimer, P. K., 157
Werner, L. C., 32
Wheeler, H. A., 589
Wheeler, M. S., 357
White, S. Y., 308
Wiegand, R. V., 153
Williams, F. C., 572
Willis, C. H., 247
Winter, W. F., 100
Woodyard, J. R., 525, 533
Wright, S. B., 313

Z

Zworykin, V. K., 152–155

SUBJECT INDEX

A

Abbreviations, 612
Acorn tubes, 89
Activation of emitting surface, oxide-coated, 9
  thoriated-tungsten, 10
Admittance, input, 260
Air-cooled tubes, 31
Air gap in chokes, 192
Amplification of alternating currents, 58
Amplification factor, 601
  of pentodes, 79
  of tetrodes, 77
  of triodes, 57, 67
Amplifiers, analysis of, 60
  balanced, d-c, 309
  buffer, 460, 483
  cathode-coupled, 458
  cathode-follower, 371
  class A, 257, 317, 602
  class AB, 257, 352, 602
  class B, 257, 395, 430, 602
    load circuit of, 397
    push-pull, 348
  class C, 257, 395, 602
    adjustment of, 427
    design of, 405, 414, 421
    grid-modulated, 521
      high-efficiency, 525
    load circuit for, 397
    neutralized, 423
    plate-modulated, 514
    power required to drive, 403, 409
    sources of power for, 422
    using pentodes and tetrodes, 413
    wave shapes of currents and emfs in, 398–403
  classification of, 257
  d-c, 257, 307
Amplifiers, distortion in (*see* Distortion)
  Doherty high-efficiency, 432
  equivalent circuit of triode, 64
  feedback in (*see* Feedback)
  frequency response of, 272, 284, 299, 357, 364
  gain of, 268, 297, 362, 372
  graphical determination of distortion in, 328, 334
  grounded-grid, 436
  high-frequency compensation of, 290
  impedance-coupled, 303, 383
  input admittance of, 260
  linear, 257, 395
  Loftin-White, 308
  low-frequency compensation of, 286
  maximum undistorted output of, 319
  "motorboating" in, 376
  multistage, 285, 394
  with negative feedback, 361–375
  neutralized, 423
  noise in, 314
  phase relations in, 64
  phototube control of, 310
  plate efficiency of, class B, 354, 432
    pentodes, class A, 340
    push-pull, 354
    triodes, class A, 338
  power, 257, 316
  power output of, 333, 335, 348, 408
  push-pull, class A, 341–348
    Class AB, 352
    class B, 348, 352
  radio-frequency, 257, 382–439
    power pentode, 413
    power triode, 394–413
    voltage, 382–394

Amplifiers, with reactive load, 337, 506
  regeneration in, 375
  resistance-coupled, 264–294
  with reversed feedback, 361
  with series-peaking circuit, 293
  space-charge-grid, 89
  transformer-coupled, 294, 356
    resistance-loading in, 302
  tuned, 383, 397
    neutralization of, 423
    optimum coupling in, 390, 393
  universal amplification curves for, 274
  video-frequency, 257, 275, 286, 290
  voltage, 264, 383
Amplitude distortion (*see* Distortion)
Amplitude limiter, 597
Amplitude modulation (*see* Modulation)
Anode, 20, 601
  graphite, 20, 115
Anode fall space, 98
Arc back, 99, 109
Arc discharge, 93, 98
Argon gas in thyratrons, 124
Aston dark space, 98
Atoms, excited state, 95
  metastable state, 95
Audio-frequency amplifier, 256
Audio-frequency transformers, 300–302
  equivalent circuit of, 295
  frequency response of, 299
  resistance loading of, 302
Autodyne detector, 586
Automatic volume control (a.v.c.) 82, 587

B

Balance coil in rectifiers, 172
Balanced modulator, amplitude modulation, 533
  frequency modulation, 549
Band-pass filter, 388
Band width, in frequency modulation, 540
  in radio-frequency amplifiers, 386, 388, 393
  in resistance-coupled amplifiers, 276
Battery charging, 117
Beam, electron, 146
Beam power tube, 82, 331
Beat detector, 578
Beat-frequency oscillator, 480
Bessel's functions, 539
Bias (*see* Grid-bias voltage)
Blocking condenser, 398, 421, 450, 601
Breakdown potential, 96
Bridge-stabilized oscillator, 475
Buffer amplifier, 460, 483

C

Candle power, 132
Capacitance, equivalent input, of a tube, 262
Carbonization of thoriated tungsten filaments, 10
Carrier current, 510
Carrier removal, 533
Carrier wave, 510, 602
  suppression of, 533
Cathode, 27, 602
  cold, 117
  disintegration of, by positive-ion bombardment, 11, 28, 99
  heat-shielded, 104
  indirectly-heated, 13, 28
  oxide-coated, 9
  thoriated-tungsten, 10
  tungsten, 8
  virtual, 83, 85, 89, 479
    (*See also* Filament)
Cathode bias, 278, 422
Cathode by-pass condenser, 284
Cathode-coupled amplifier, 438
Cathode dark space, 98
Cathode-follower amplifier, 371
Cathode glow, 98

Cathode grid, 77
Cathode heating power, 15
Cathode pool, 105
Cathode-ray tube, 145, 303
  for checking modulation, 528
Cathode spot, 108
Cesiated silver photoemitter, 20, 136
Characteristic curves, composite, 344
  dynamic, tetrode, 74
    triode, 59, 318
  static, beam tube, 85
    diode, 37
    pentode, 78
    phototube, 133, 135
    tetrode, 72, 130
    thyratron, 123
    triode, 51, 123
Characteristic surface, 53
Child's equation, 44
Choke, design of, for rectifier service, 191
  swinging, 193
Class A, AB, B, and C amplifiers (see Amplifiers)
Clipping of signal wave in detectors, 578, 589
Coefficient of coupling, 390, 393
Cold-cathode tubes, 117, 225
  grid-controlled, 126, 249
Color response, of photoconductive cells, 140
  of photosensitive surfaces, 21
  of phototubes, 136
Colpitts oscillator, 444
Commutating condenser, 241
Condenser-input filter, 181, 199
Control grid, 77, 603
Constant-voltage regulator, 225
Control device, vacuum tube as, 217–254
Controlled multivibrators, 482
Controlled rectifier, 236
Conversion transconductance, 580, 603
Converter, pentagrid, 88, 583
Cooling of vacuum tubes, 29
Copper oxide rectifier, 210
Coupling condenser, 277
Crookes dark space, 98
Cross modulation, 531
Crystal, quartz, 468–475
    AT, BT, CT, DT cuts, 471
    equivalent circuit of, 473
    GT cut, 472
    X and Y cuts, 469
Crystal oscillator, 467
Current regulator, 226
Cutoff grid voltage, 55, 349, 395, 603
Cyclophon, 559, 599

D

Decibel, 259
Decoupling circuits, 377
Definitions of vacuum-tube terms, 601–607
Degeneration, 375
Deionization time, 124, 228
Demodulation defined, 563
Demodulators (see Detectors)
Demountable tubes, 31
Detection plate resistance, 569
Detectors, 563
  autodyne, 586
  filters for, 564, 588
  frequency-modulation, 595
  grid-leak, 577
  heterodyne, 555, 557, 578
  infinite-impedance, 573
  oscillating, 586
  peak linear, 564
  pulse-modulation, 598
  regenerative, 587
  square-law, 575
Diodes, demodulators, 563, 574
  gas-filled, 100
  high-vacuum, 36
  modulators, 527
  plate resistance of, 47
Direct-current amplifiers, 257, 307
Direct-current transmission of power, 217, 248
Discriminator, 595

Disintegration of cathodes, 11, 99
Distortion, amplitude, 258
analysis of, 495
with a v c., 589
clipping, 573, 589
components produced by, 605
with constant excitation voltage, 328
definition of, 258
due to grid current, 318, 351
effect of load resistance on, 327
equation of, for pentode or beam amplifier tube, 335
for triode, 333
with feedback, 363
frequency, 258, 272, 299, 357
graphical determination of, 332, 335, 347, 352
in grid-modulated class C amplifier, 523
harmonic, 258
with maximum power output, 322
in peak linear detector, 570
phase, 259
in plate-modulated class C amplifier, 519
in single-side-band transmission, 535
in square-law detector, 576
Doherty high-efficiency amplifier, 432
Double ionization, 100
Doubler, frequency, 429, 552
voltage, half-wave, 205
Driving power for class C amplifier, 403, 410
Dumet, 25
Duo-diode-pentode, 89
Duo-diode-triode, 88
Duplex tubes, 88
Dynamic characteristic curves (*see* Characteristic curves)
Dynatron, oscillator, 477
in parasitic oscillations, 486

E

Edison effect, 5
Efficiency of an amplifier (*see* Plate efficiency)
Electron beam, 83, 146, 158, 560
Electron coupling, 460, 584
Electron gun, 145, 155, 559
Electron multiplier, 151
Electron optical system, 146
Electron-ray tube, 150
Electron theory, 5
Electrons, charge of, 5
emission velocity of, 8, 19
free, 6
high-field emission of, 17
mass of, 5
mean free path of, 34, 49
photoelectric emission of, 18
properties of, 5
secondary emission of, 16, 148, 477, 560
thermionic emission of, 7
transit time of, 37, 54, 91, 259
Emission, high-field, 17
laws governing, 13
measurement of, 11
secondary, 16, 73, 148, 477, 560
thermionic, 7–15
Emission efficiency, 9, 104
Emission velocity of electrons, 8, 19, 575
Emitters, electron (*see* Cathode, Filament)
Energy storage, amplifiers, 399, 416
Equivalent circuit, crystal, 473
dynatron, 478
photovoltaic cell, 143
triode amplifier, 64
Evacuation of vacuum tubes, 33
Excitation voltage, 603

F

Faraday dark space, 93
Feed-back factor, 363
Feedback, in audio amplifiers, 361
circuits for, 368
current, 367
effect of, on frequency response, 364
in grid-modulated amplifiers, 521
in radio amplifiers, 424

Feedback, in audio amplifiers, voltage, 361
(*See also* Regeneration)
Filament, effect of voltage drop in, 45
(*See also* Cathode)
Filters, band-pass, 388
for detector output, 564, 588
for power rectifiers, 181–193, 197
L-section, 184–191
pi-section, 183, 190
tuned-circuit, 186
resistance-capacitance, 377
Fixed bias, 397, 422
Fleming valve, 37
Fluorescent screen, 145, 149
Foot-candle, 132
Fourier analysis, 329, 613, 617
Free-grid potential, 54, 68
Frequency of oscillation, of a crystal, 474
of an R-C oscillator, 465
of a tuned-circuit oscillator, 452
Frequency control with resonant lines, 476
Frequency demodulation, 595
Frequency distortion (*see* Distortion)
Frequency doublers, 429, 552
Frequency modulation, balanced-modulator system, 549
definition, 509
equations of, 537
by reactance tube, 554
synchronized, 556
using eighth-wave line, 548
vector diagrams of, 545–547
Frequency stability, of crystal oscillators, 474
of tuned-circuit oscillators, 457
Frequency tolerance, 468

G

Gas, current flow through, 93
in vacuum tubes, 43, 93–131
removal of, 33
Gas-filled tubes, 23, 93–131, 135
Getter, 34, 115

Glow discharge, 93, 97
Glow-discharge tube, 225
Graphite anode, 29, 115
Grid, action of, 51, 54–57, 119–122
cathode, 77
in cold-cathode tube, 126
construction of, 28
control, 77
cutoff potential of, 55
effect of driving positive, 318, 350, 523
free, 54, 68
power required to drive, 403, 410
screen, 70, 77, 605
space-charge, 89, 606
suppressor, 77, 524, 606
thyratron, 122
Grid-bias voltage, 165, 422, 603
by grid leak, 422, 444
for plate-modulated class C amplifier, 520
by self-bias, 278, 422
Grid condenser in oscillators, 461
Grid-controlled cold-cathode tube, 126
Grid current, 68, 130
effect of, on distortion, 70, 318, 350
Grid leak, 265, 277, 422, 444, 577
Grid modulation, 521–525
Grid phase-shift control, of ignitrons, 246
of thyratrons, 231
Grounded-grid amplifier, 436

H

Harmonics, 603
in class A amplifiers, 326, 334
in class B and C amplifiers, 396, 419, 617
even, cancellation, in push-pull amplifiers, 342
(*See also* Distortion)
Hartley oscillator, 443
Hazeltine system of neutralization, 425
Heat-shielded cathode, 104
Heater, 13, 604

Heising system of modulation, 516
Heterodyne, 555, 557, 578
High-vacuum tube, 23, 26-92
Hum, 315, 342, 348, 363

I

Iconoscope, 155
Ignitron, 112
  firing circuits for, 206
  inverter, 247
  phase-shift control of, 246
Image orthicon, 157
Impedance-coupled amplifiers (*see* Amplifiers, impedance-coupled)
Incremental permeability, 192
Inductance-coupled amplifiers (*see* Amplifiers, impedance-coupled)
Inductance-input filter, 181
Infinite-impedance detector, 573
Intermediate frequency, superheterodyne, 582
Intermittent oscillations, 462, 487
Interphase reactor, 172
Inverse voltage, 90, 106, 619
Inverters, mercury-arc, 245
  thyratron, self-excited, 243
    separately excited, 240
Ionization, 48, 95-100
  double, 100
  in phototubes, 135
Ionization potential, 96
Ions, 7, 94, 102, 116, 120
Iron selenide cell, 142

K

Keep-alive anodes, 107
Keep alive circuit, 212
Kinescope, 153
Konal, 13

L

Lead sulfide cell, 140
Lighthouse tubes, 90, 477
Lighting control, saturable reactors, 236, 247, 248
Limiter, 312, 550, 597
Linear amplifier, class B, 257, 305
Linear detector, 563-573
Linear modulator, 514
Lissajous figures, 307
Load line, 322, 347
Loftin-White amplifier, 308
L-section filter, 181
  design of, 187
Lumen, 132

M

Master - oscillator - power - amplifier (MOPA), 460
Maximum positive value of grid voltage, class C amplifier, 398
Mean free path of an electron, 31, 49
Mercury-arc rectifier, 211
Mercury-arc tube, 106
Mercury-pool cathode, 106
Mercury-vapor tube, 100
Microphonic noise, 316
Minimum plate voltage, in a class A amplifier, 340
  in a class C amplifier, 398, 428
Modulated amplifier, grid-modulated, 521
  plate-modulated, 514
  square-law, 526
  suppressor-grid, 524
Modulated wave, 511, 604
  analysis of, amplitude, 511
    frequency and phase, 535
  power in, 513
  pulse, 558
Modulating wave, 510
Modulation, 508
  amplitude, 509, 510
  checking with cathode-ray tube, 525
  classification of, 509
  components of amplitude, 531
  components of frequency, 539
  frequency, 509, 536-557
  grid (class C), 521, 525
  phase, 509, 536
  plate, 514-520

Modulation, pulse, 557
  square-law, 526–533
  suppressor-grid, 524
Modulation factor, 511, 530
Modulation index, 539
Modulators, 508
  balanced, 533, 549
  diode, 527
  power requirements of plate-modulated, 517
  push-pull, 534
  reactance-tube, 553
  square-law (van der Bijl), 528
Motor, thyratron, 247, 248
"Motorboating" in amplifiers, 376
Mu factor, 280, 414, 604
Multigrid tubes, 70–91
Multivibrators, 481, 559, 598
  controlled, 482
Mutual conductance (see Transconductance)

N

Negative feed-back amplifier, 351–375
Negative glow region, 98
Negative-grid-control thyratron, 124
Negative resistance in tetrodes, 76
Negative transconductance oscillator, 479
Neutralized amplifiers, 423
Noise in amplifiers, 314
Noise reduction, due to feedback, 368
  in frequency modulation, 546
Nomenclature, 607–611
Norton's theorem, 267

O

Orthicon, 157
Oscillations, parasitic, 484
Oscillators, 441–487
  beat-frequency, 480
  bridge-stabilized, 475
  circuits of, 443
Oscillators, Colpitts, 444
  crystal, 467
  design of, 454
  dynatron, 477
  electron-coupled, 460
  frequency of, 452, 465, 474
  frequency-controlling, 454, 467
  frequency stability of, 457, 474
  Hartley, 443
  intermittent operation of, 461, 487
  negative transconductance, 479
  phase correction of, 459
  phase-shift, 467
  power, 454
  R-C, 462
  relaxation, 304, 481
  sweep-circuit, 304
  tuned-grid, tuned-plate, 445
  vector diagram of, 458
Oscillatory circuit, general equation of, 445
Oscillograph, cathode-ray, 304
Oscilloscope, 149, 303
  linear time axis for, 305
Output maximum undistorted, of an amplifier, 319
Oxide-coated cathodes, 9, 12

P

Parallel feed, 398, 401
Parasitic oscillations, 484
Peak linear detector, 564
Peak tube current, 194, 621
Peak voltmeters, 593
Pentagrid converter, 88, 583
Pentagrid mixer tube, 88, 581
Pentode tubes, 77
Phase correction to improve frequency stability, 459
Phase distortion (see Distortion)
Phase-inverting tube, 355
Phase modulation, definition of, 509
  equations of, 537
Phase shift in feed-back amplifiers, 364

Phase-shift control, of ignitrons, 246
of thyratrons, 231
Phase-shift oscillator, 467
Photocathode, 133, 155
Photoconductive cells, 138
Photoelectric emission, 18–22
Photoemissive tube, 132
Photon, 19
Phototubes, 132
circuits for, 138, 251, 310
Photovoltaic cells, 141, 251
Photronic cell, 142
Pi-section filter, 181, 199
Piezoelectric, 468
Planck's constant, 19
Plasma, 98, 102
Plate current, power-series expansion of, diodes, 497
triodes and pentodes, 488
Plate efficiency, beam tubes, 340
class B, 355, 432
class C, 394, 408
pentodes, class A, 340
push-pull amplifier, 354
triodes, class A, 338
Plate loss, 45
Plate modulation, 514–520
Plate resistance, a-c, 47, 61, 67
detection, 569
of pentode tubes, 79
of screen-grid tubes, 76
Polarization selectivity, 21
Positive-grid-control thyratron, 121
Positive-ion bombardment, 11, 28, 99
Positive-ion sheath, 98, 103
Potential barrier, 6, 19
Potential distribution, in beam tubes, 85
in diodes, high-vacuum, 39
in mercury-vapor tubes, 101
in pentodes, 79
in tetrodes, 75
in triodes, high-vacuum, 54
gas-filled, (thyratrons), 121
Power, in a modulated wave, 513
transmission of, by direct current, 247, 248
Power amplifiers, 257, 316–361, 394–439
Power emission chart, 15
Power oscillators, 454
Power output of amplifiers, 333, 335, 408
Power-series equation of plate current, 488, 497
Pulse demodulation, 598
Pulse modulation, 557
Push-pull amplifiers, class A, 341
class AB, 352
class B, 348, 352

Q

Q, of coils, 384, 393
of crystals, 473
of tank circuits, 415, 419, 454
Quantum theory, 19
Quartz crystals (see Crystal)

R

Radio-frequency amplifiers (see Amplifiers)
R-C oscillators, 462
Reactance-tube modulator, 553
Reactivation of thoriated-tungsten filaments, 12
Reactive loads for amplifiers, 337, 506
Receivers, superheterodyne, 582
Rectification characteristic curves, 567
Rectification factor, 569
Rectifiers, circuits for, 164–177
controlled, 211, 236
copper oxide, 210
design of, 187, 196
filters for, 181, 199
ignitron, 112, 206
mercury-arc, 106, 211
using r-f power, 202
table of design data for, 195, 618
tungar, 115
voltage-multiplying, 204
voltage regulation of, 224

Regeneration, in amplifiers, 375
in detectors, 587
Regulator, 224
Relaxation oscillations, 304, 481, 487
Relays, control of, 221
Remote cutoff tubes, 82
Resistance, a-c plate, 47, 61, 67
negative, 76, 477
Resistance-capacitance filter, 377
Resistance-coupled amplifiers (*see* Amplifiers)
Resistance loading of transformers, 302
Resonant lines as frequency-controlling elements, 476
Reversed-feedback amplifiers, 361
Rice method of neutralization, 426
Richardson's work function, 14
Rochelle salts, 468

S

Saturated-reactor method of thyratron control, 235
Saturation of iron cores, 192, 235, 299–300, 342, 518
Saturation current, 39, 134
Screen grid, 77, 605
Screen-grid by-pass condenser, effect of, 278
Screen-grid tubes, 70
gas-filled, 129
Secondary emission, 16
in an electron multiplier, 151
in screen-grid tubes, 73, 78
Selectivity, 389, 582
Selenium cell, 138
Series feed, 398
Series-peaking circuit, 293
Sheaths, positive-ion, 98, 102
Shielding of tubes, 25, 71
Shot effect, 314
Side bands, 512, 539, 606
power in, 513
Single side-band transmission, 534
Skin effect in iron, 383
Space charge, 39, 52, 101, 153, 314
Space-charge equation, 43
Space-charge-grid tubes, 89
Space current in screen-grid tubes, 72
Spectral selectivity, 21
Square-law demodulation, components of, 577
Square-law detector, 563, 575
Square-law modulator, 526
components in, 531
performance of, 529
Starting voltage, 120
Static characteristic curves (*see* Characteristic curves)
Superheterodyne receiver, 582–586
Suppressor grid, 77, 81, 606
Suppressor-grid modulation, 524
Sweep-circuit oscillator, 304
Swinging choke, 193
Symbols, 607–611
Synchronized frequency modulation, 556

T

Tank circuit, 398, 600
approximate solution for, 421
design of, 414–422
Taylor's theorem, 489, 499
Television, tubes for, 153
Temperature limits of mercury-vapor tubes, 105
Tetrodes, 70
Thallous sulfide cell, 140
Thermal agitation, 314
Thermionic emission, 7–15
Thoriated-tungsten cathodes, 10, 14
carbonization of, 10
Threshold frequency for photoelectric emission, 20
Thyratron inverters, 239–245
Thyratron motor, 247, 248
Thyratrons, construction of, 120
d-c circuits for, 227
grid phase shift of, 231
phototube control of, 252
potential distribution curves of, 121
screen-grid, 129
starting voltage of, 120
Tourmaline, 468

Transconductance, 61, 67, 76, 79, 275, 606
Transformer-coupled amplifiers (*see* Amplifiers)
Transit time of the electron, 37, 54, 91, 259
Transitron oscillator, 479
Transmission line, as frequency controlling element, 476
  as frequency modulator, 548
Transrectification characteristic curves, 567
Transrectification factor, 569
Triodes, 50–70, 119–129
  acorn-type, 89
  characteristics of, 51–53
  equivalent input capacitance of, 262
  gas-filled, 119
Tubes, construction details of, 23–33
Tuned-grid-tuned-plate oscillator, 445
Tungar rectifier, 115
Tungsten cathodes, 8

U

Universal amplification curves, 274
Utilization factor, 198, 621, 625

V

Vacuum-tube characteristics, measurement of, 37, 51, 567
Vacuum-tube coefficients, 67, 76
Vacuum-tube nomenclature, 607–611
Vacuum-tube voltmeters, 591–595
Vacuum tubes, construction of, 23–33
  evacuation of, 33
Van der Bijl modulator, 523
Variable-mu tubes, 82
Vector diagram, of amplifiers, 64–67
  of frequency modulation, 545
  of input circuit to a tube, 263
  of oscillators, 458
Video-frequency amplifiers, 257, 275, 286, 290
Virtual cathode, 83, 85, 89, 479
Voltage amplifiers, a-f, 257, 261–316
  r-f, 257, 383–394
Voltage-multiplying circuits, 204
Voltage regulator, 110, 221, 217
Voltmeter, vacuum-tube, 591–595
Volume expanding, 312
Volume limiting, 312
Volume compressing, 312
Volume-control methods, 311

W

Water-cooled tubes, 30, 110, 113
Welding, resistance, 217, 218
Work function, 8, 14, 20

X

X-ray tubes, 153

Z

Zero-grid-bias tubes, 351
Zero-temperature-coefficient crystals, 470